

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:37:16 ; Search time 38.3333 Seconds

(without alignments)  
122.745 Million cell updates/sec

Title: US-10-040-128-1

Perfect score: 86  
Sequence: 1 DEINQMHTDEKNE 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubppa/PCUUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/prodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	15	US-10-040-128-1	Sequence 1, Appli
2	78	90.7	15	US-10-040-128-2	Sequence 2, Appli
3	73	84.9	15	US-10-040-128-3	Sequence 3, Appli
4	73	84.9	108	US-10-006-863-6	Sequence 6, Appli
5	73	84.9	108	US-10-395-032-6	Sequence 6, Appli
6	73	84.9	784	US-09-905-983-52	Sequence 52, Appli
7	73	84.9	784	US-10-211-462-105	Sequence 105, Appli
8	73	84.9	784	US-10-021-660-99	Sequence 99, Appli
9	73	84.9	784	US-10-174-677-23	Sequence 23, Appli
10	46	53.5	95	US-10-424-599-281094	Sequence 281094,
11	46	53.5	183	US-09-925-298-550	Sequence 530, App
12	46	53.5	183	US-10-102-806-550	Sequence 530, App
13	46	53.5	192	US-10-188-186-28	Sequence 28, Appli
14	46	53.5	193	US-09-764-853-673	Sequence 673, App
15	45	52.3	418	US-10-282-122A-74566	Sequence 74566, A

16	45	52.3	1312	US-10-369-493-1904	Sequence 1904, Ap
17	44	51.2	88	US-10-424-599-186878	Sequence 186878,
18	44	51.2	165	US-10-424-599-177622	Sequence 177622,
19	44	51.2	387	US-10-424-599-233597	Sequence 233597,
20	44	51.2	437	US-10-425-114-39166	Sequence 39166, A
21	44	51.2	558	US-10-369-493-6392	Sequence 6392, Ap
22	43	50.0	731	US-10-424-599-172399	Sequence 172399,
23	43	50.0	105	US-09-864-408A-6804	Sequence 6804, Ap
24	43	50.0	367	US-10-282-122A-51530	Sequence 51530, A
25	42	48.8	166	US-10-437-963-142823	Sequence 142823,
26	42	48.8	230	US-10-437-963-104838	Sequence 104838,
27	42	48.8	254	US-09-864-761-33770	Sequence 33770, A
28	42	48.8	261	US-10-424-599-283374	Sequence 283374,
29	42	48.8	515	US-10-424-599-166030	Sequence 166030,
30	42	48.8	1039	US-09-900-237-14	Sequence 14, Appli
31	42	48.8	1103	US-10-408-765A-2255	Sequence 2255, Ap
32	41	47.7	122	US-10-424-599-187689	Sequence 187689,
33	41	47.7	167	US-10-243-552-596	Sequence 596, App
34	41	47.7	294	US-10-424-599-153982	Sequence 153982,
35	41	47.7	494	US-10-369-493-20846	Sequence 20846, A
36	41	47.7	549	US-10-424-599-186879	Sequence 186879,
37	41	47.7	551	US-10-424-599-186880	Sequence 186880,
38	41	47.7	587	US-10-424-599-235268	Sequence 235268,
39	41	47.7	657	US-10-276-774-1900	Sequence 1900, Ap
40	41	47.7	712	US-09-978-295A-264	Sequence 264, App
41	41	47.7	722	US-09-978-697-264	Sequence 264, App
42	41	47.7	722	US-09-978-192A-264	Sequence 264, App
43	41	47.7	722	US-09-999-832A-264	Sequence 264, App
44	41	47.7	722	US-09-978-189-264	Sequence 264, App
45	41	47.7	772	US-09-978-608A-264	Sequence 264, App

#### ALIGNMENTS

RESULT 1  
US-10-040-128-1  
Sequence 1, Application US/10040128  
Publication No. US20020160003A1  
GENERAL INFORMATION:  
APPLICANT: Liao, Fang  
APPLICANT: Hicklin, Daniel J.  
APPLICANT: Bohlen, Peter  
TITLE OF INVENTION: Antibody Antagonists of VE-Cadherin Without Adverse Effects on  
FILE REFERENCE: 11245/46976  
CURRENT APPLICATION NUMBER: US/10/040,128  
CURRENT FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: WordPerfect 8.0 for Windows  
SEQ ID NO 1  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: synthetic peptide  
US-10-040-128-1

Query Match 100.0%; Score 86; DB 13; Length 15;  
Best Local Similarity 100.0%; Pred. 1.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DEINQMHTDEKNE 15  
| | | | | | | | | | | | | | | | |  
Db 1 DEINQMHTDEKNE 15

RESULT 2  
US-10-040-128-2  
Sequence 2, Application US/10040128  
Publication No. US20020160003A1  
GENERAL INFORMATION:  
APPLICANT: Liao, Fang

```
; APPLICANT: Hicklin, Daniel J.
; APPLICANT: Bohlen, Peter
; TITLE OF INVENTION: Antibody Antagonists of Vb-Cadherin Without Adverse Effects on
; FILE REFERENCE: 11245/46976
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; US-10-040-128-2

Query Match          90.7%; Score 78; DB 13; Length 15;
Best Local Similarity 93.3%; Pred. No. 2e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEINQMHIIDEKN 15
Db      1 DMIWMQMHIIDEKN 15

RESULT 3
US-10-040-128-3
; Sequence 3, Application US/10040128
; Publication No. US2002016003A1
; GENERAL INFORMATION:
; APPLICANT: Liao, Fang
; APPLICANT: Hicklin, Daniel J.
; APPLICANT: Bohlen, Peter
; TITLE OF INVENTION: Antibody Antagonists of Vb-Cadherin Without Adverse Effects on
; FILE REFERENCE: 11245/46976
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; US-10-040-128-3

Query Match          84.9%; Score 73; DB 13; Length 15;
Best Local Similarity 92.9%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEINQMHIIDEKN 14
Db      1 DMIWMQMHIIDEKN 14

RESULT 4
US-10-006-869-6
; Sequence 6, Application US/10006869
; Publication No. US2003008216A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gout, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C7
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentln Ver. 2.0
```

```
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-006-869-6

Query Match          84.9%; Score 73; DB 14; Length 108;
Best Local Similarity 92.9%; Pred. No. 0.00092;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEINQMHIIDEKN 14
Db      5 DMIWMQMHIIDEKN 18

RESULT 5
US-10-395-032-6
; Sequence 6, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gout, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C9
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-395-032-6

Query Match          84.9%; Score 73; DB 15; Length 108;
Best Local Similarity 92.9%; Pred. No. 0.00092;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEINQMHIIDEKN 14
Db      5 DMIWMQMHIIDEKN 18

RESULT 6
US-09-905-983-52
; Sequence 52, Application US/09905983
; Patent No. US20020045591A1
; GENERAL INFORMATION:
; APPLICANT: Geiger, Benjamin
; APPLICANT: Ben-Ze'ev, Avri
; APPLICANT: Sadot, Eilat
; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER
; FILE REFERENCE: 01/22326
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 52
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-905-983-52

Query Match          84.9%; Score 73; DB 9; Length 784;
Best Local Similarity 92.9%; Pred. No. 0.0072;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEINQMHIIDEKN 14
Db      48 DMIWMQMHIIDEKN 61
```

```

RESULT 7
US-10-211-462-105
; Sequence 105, Application US/10211462
; Publication No. US2004003495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Atiz, Natasia
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-105

```

```

Query Match      84.9%; Score 73; DB 12; Length 784;
Best Local Similarity 92.9%; Pred. No. 0.0072;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 DEINQMHIDEKN 14
      |||||
Db      48 DWNNHMHDOQ 61

```

```

RESULT 8
US-10-021-660-99
; Sequence 99, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis,
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-99

```

```

Query Match      84.9%; Score 73; DB 14; Length 784;
Best Local Similarity 92.9%; Pred. No. 0.0072;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 DEINQMHIDEKN 14

```

```

Db      48 DWNNHMHDOQ 61
      |||||
RESULT 9
US-10-174-677-23
; Sequence 23, Application US/10174677
; Publication No. US2003019070A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR
; FILE REFERENCE: 40716(1P-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-677-23

```

```

Query Match      84.9%; Score 73; DB 14; Length 784;
Best Local Similarity 92.9%; Pred. No. 0.0072;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 DEINQMHIDEKN 14
      |||||
Db      48 DWNNHMHDOQ 61

```

```

RESULT 10
US-10-424-599-281094
; Sequence 281094, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 281094
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9584C.1.pap
US-10-424-599-281094

```

```

Query Match      53.5%; Score 46; DB 12; Length 95;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 DEINQMHIDEK 12
      |||||
Db      18 DQNNHMHDOQ 29

```

```

RESULT 11
US-09-925-298-530
; Sequence 530, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10

```

```

; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 530
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (79)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-530

Query Match      53.5%; Score 46; DB 12; Length 183;
Best Local Similarity 40.0%; Pred. No. 25;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy      1 DEINOMHIDEKNE 15
Db      122 EEWMDKLMMDKEKE 136

RESULT 12
US-10-102-806-530
; Sequence 530; Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: 2002-03-22
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 530
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (79)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-530

Query Match      53.5%; Score 46; DB 14; Length 183;
Best Local Similarity 40.0%; Pred. No. 25;
```

```

Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy      1 DEINOMHIDEKNE 15
Db      122 EEWMDKLMMDKEKE 136

RESULT 13
US-10-188-186-28
; Sequence 28; Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 28
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-28

Query Match      53.5%; Score 46; DB 12; Length 192;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy      1 DEINOMHIDEKNE 15
Db      131 EEWMDKLMMDKEKE 145

RESULT 14
US-09-764-853-673
; Sequence 673; Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 673
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```



```

; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (103)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-673

```

```

Query Match      53.5%; Score 46; DB 9; Length 193;
Best Local Similarity 40.0%; Pred. No. 26;
Matches      6; Conservative      7; Mismatches      2; Indels      0; Gaps      0;

```

```

Qy      1 DEINOMHIDEKNE 15
Db      132 EEVWDKIMDKERKE 146

```

```

RESULT 15
US-10-282-122A-74566
; Sequence 74566, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO 74566
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74566

```

```

Query Match      52.3%; Score 45; DB 12; Length 418;
Best Local Similarity 42.9%; Pred. No. 85;
Matches      6; Conservative      5; Mismatches      3; Indels      0; Gaps      0;

```

```

Qy      2 EIMNOMHIDEKNE 15
Db      14 EIMDAITAEERKE 27

```

Search completed: July 29, 2004, 11:50:32  
Job time : 38.3333 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:32:25 ; Search time 13 Seconds

(without alignments)  
59,568 Million cell updates/sec

Title: US-10-040-128-1

Sequence: 1 DEINWQHIDEKNE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/ECTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	84.9	108	4	US-09-187-859-6
2	73	84.9	108	4	US-09-839-542B-6
3	73	84.9	108	4	US-09-535-852-6
4	73	84.9	780	1	US-08-188-228-50
5	73	84.9	780	1	US-08-332-643-44
6	73	84.9	780	1	US-08-332-638-50
7	40	46.5	747	4	US-09-134-000C-6641
8	40	46.5	1114	4	US-09-637-145-4
9	40	46.5	1122	4	US-10-072-094-8
10	40	46.5	1122	4	US-10-072-094-91
11	39.5	45.9	138	4	US-09-570-921-5
12	39.5	45.9	198	3	US-08-965-036-87
13	39.5	45.9	613	3	US-09-257-490-14
14	39	45.3	148	4	US-09-621-976-4783
15	39	45.3	226	4	US-09-533-029-88
16	39	45.3	231	4	US-09-345-473E-36
17	39	45.3	535	2	US-08-933-750C-20
18	39	45.3	535	3	US-09-234-613-20
19	39	45.3	535	4	US-09-023-942A-10
20	38.5	44.8	138	4	US-09-570-921-6
21	38.5	44.8	198	3	US-08-965-056-85
22	38.5	44.8	417	1	US-08-118-469A-5
23	38.5	44.8	417	1	US-08-909-119-5
24	38	44.2	110	4	US-09-187-859-13
25	38	44.2	110	4	US-09-839-542B-13
26	38	44.2	110	4	US-09-535-852-13
27	38	44.2	135	4	US-09-732-210-121

28	38	44.2	163	4	US-10-072-094-4	Sequence 4, Appli
29	38	44.2	326	4	US-09-345-473E-10	Sequence 10, Appl
30	38	44.2	532	1	US-08-188-228-44	Sequence 44, Appl
31	38	44.2	532	1	US-08-332-638-44	Sequence 44, Appl
32	38	44.2	653	4	US-09-107-532A-4502	Sequence 4502, Ap
33	38	44.2	734	4	US-09-134-000C-5602	Sequence 5602, Ap
34	38	44.2	780	4	US-10-072-094-93	Sequence 93, Appl
35	38	44.2	793	1	US-08-332-643-48	Sequence 54, Appl
36	38	44.2	793	1	US-08-332-638-54	Sequence 48, Appl
37	38	44.2	793	1	US-08-332-638-54	Sequence 54, Appl
38	38	44.2	799	1	US-08-188-228-42	Sequence 42, Appl
39	38	44.2	799	1	US-08-332-638-42	Sequence 90, Appl
40	38	44.2	800	4	US-09-555-790A-2	Sequence 2, Appli
41	38	44.2	855	4	US-10-072-094-9	Sequence 9, Appli
42	38	44.2	879	4	US-10-072-094-90	Sequence 2, Appli
43	38	44.2	938	4	US-09-637-145-2	Sequence 89, Appl
44	38	44.2	1011	4	US-10-072-094-89	Sequence 87, Appl
45	38	44.2	1069	4	US-10-072-094-87	

#### ALIGNMENTS

```

RESULT 1
US-09-187-859-6
; Sequence 6, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-187-859-6

Query Match      84.9%; Score 73; DB 4; Length 108;
Best Local Similarity 92.9%; Pred. No. 0.00025;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEINWQHIDEKKN 14
Db      5 DWINWQHIDEKKN 18

RESULT 2
US-09-839-542B-6
; Sequence 6, Application US/09839542B
; Patent No. 6569996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-542B-6

```

Query Match 84.9%; Score 73; DB 4; Length 108;  
Best Local Similarity 92.9%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 14  
| | | | | | | | | | | | | | | |  
Db 5 DWINWQHIDEKN 18

RESULT 3  
US-09-535-852-6

Sequence 6, Application US/09535852

Patent No. 6638911

GENERAL INFORMATION:

APPLICANT: Blachuk, Orest W.

APPLICANT: Symonds, James M.

APPLICANT: Gour, Barbara J.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING

FILE REFERENCE: 100086.407C6

CURRENT APPLICATION NUMBER: US/09/535,852

CURRENT FILING DATE: 2001-05-21

NUMBER OF SEQ ID NOS: 2009

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 108

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-535-852-6

Query Match 84.9%; Score 73; DB 4; Length 108;  
Best Local Similarity 92.9%; Pred. No. 0.00025;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 14  
| | | | | | | | | | | | | | | |  
Db 5 DWINWQHIDEKN 18

RESULT 4  
US-08-188-228-50

Sequence 50, Application US/08188228

Patent No. 5597725

GENERAL INFORMATION:

APPLICANT: Suzuki, Shintaro

TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESS:

ADDRESSER: Marshall, O'Toole, Gerstein, Murray &

STREET: 6300 Sears Tower, 233 S. Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: PC floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/188,228

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,460

FILING DATE: 19 APR 1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/872,643

FILING DATE: 17 APR 1992

ATTORNEY/AGENT INFORMATION:

NAME: No. 5597725and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31340

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 780 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-188-228-50

Query Match 84.9%; Score 73; DB 1; Length 780;  
Best Local Similarity 92.9%; Pred. No. 0.0022;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 14  
| | | | | | | | | | | | | | | |  
Db 44 DWINWQHIDEKN 57

RESULT 5  
US-08-332-643-44

Sequence 44, Application US/08332643

Patent No. 5639634

GENERAL INFORMATION:

APPLICANT: Suzuki, Shintaro

TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSER: Marshall, O'Toole, Gerstein, Murray &

STREET: Two First National Plaza, 20 South Clark

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: PC floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/332,643

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/872,643

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: No. 5639634and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/30795

TELEPHONE: (312) 346-5750

TELEFAX: (312) 984-9740

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 780 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-332-643-44

Query Match 84.9%; Score 73; DB 1; Length 780;  
Best Local Similarity 92.9%; Pred. No. 0.0022;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 14  
| | | | | | | | | | | | | | | |

Db 44 DWINOMHIDEKN 57

## RESULT 6

US-08-332-638-50  
; Sequence 50, Application US/08332638  
; Patent No. 5646250

; GENERAL INFORMATION:

; APPLICANT: Suzuki, Shinro

; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

; NUMBER OF SEQUENCES: 62

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp;

; ADDRESS: Borun

; STREET: 6300 Sears Tower, 233 S. Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/332,638

; FILING DATE: 01-NOV-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/872,643

; FILING DATE: 17 APR 1992

; APPLICATION NUMBER: US/08/049,460

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5646250and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 31340

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 50:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 780 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-332-638-50

; Query Match 84.9%; Score 73; DB 1; Length 780;

; Best Local Similarity 92.9%; Pred. No. 0.0022;

; Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

; Db 44 DWINOMHIDEKN 57

; QY 1 DEINOMHIDEKN 14

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; SEQ ID NO 6641  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6641

## Query Match

Best Local Similarity 46.5%; Score 40; DB 4; Length 747;  
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;QY 2 EWINOMHIDEKN 14  
; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; SEQ ID NO 6641  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6641

## Query Match

Best Local Similarity 46.5%; Score 40; DB 4; Length 747;  
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;QY 2 EWINOMHIDEKN 14  
; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; SEQ ID NO 6641  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6641

## Query Match

Best Local Similarity 46.5%; Score 40; DB 4; Length 747;  
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;QY 2 EWINOMHIDEKN 14  
; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; SEQ ID NO 6641  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6641

## Query Match

Best Local Similarity 46.5%; Score 40; DB 4; Length 747;  
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;QY 2 EWINOMHIDEKN 14  
; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; SEQ ID NO 6641  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6641

## Query Match

Best Local Similarity 46.5%; Score 40; DB 4; Length 747;  
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;QY 2 EWINOMHIDEKN 14  
; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

; | | | | | | | | | |

Db 789 DTWNEMH 796

RESULT 10  
US-10-072-094-91  
Sequence 91, Application US/10072094

Patent No. 6600351  
GENERAL INFORMATION:  
APPLICANT: JACKSON, DONALD  
APPLICANT: LORENZI, MATTHEW  
APPLICANT: ATTAR, RICARDO  
APPLICANT: GOTTARDIS, MARCO  
TITLE OF INVENTION: NOVEL HUMAN HISTONE PEPTIDES  
FILE REFERENCE: 3053-4145US1  
CURRENT APPLICATION NUMBER: US/10/072,094  
CURRENT FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 60/238,296  
PRIOR FILING DATE: 2001-06-14  
NUMBER OF SEQ ID NOS: 127  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 91  
LENGTH: 1122  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-072-094-91

Query Match 46.5%; Score 40; DB 4; Length 1122;  
Best Local Similarity 62.5%; Pred. No. 4.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINQM 8  
Db 789 DTWNEMH 796

RESULT 11  
US-09-570-921-5  
Sequence 5, Application US/09570921

Patent No. 6455265  
GENERAL INFORMATION:  
APPLICANT: SERIES, PIERRE-FRANCOIS  
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE  
TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION  
FILE REFERENCE: 106213  
CURRENT APPLICATION NUMBER: US/09/570,921  
CURRENT FILING DATE: 2000-05-15  
PRIOR APPLICATION NUMBER: PCT/FR98/02447  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: FR/97/14387  
PRIOR FILING DATE: 1997-11-17  
NUMBER OF SEQ ID NOS: 144  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 138  
TYPE: PRT  
ORGANISM: Human  
US-09-570-921-5

Query Match 45.9%; Score 39.5; DB 4; Length 138;  
Best Local Similarity 26.3%; Pred. No. 51;  
Matches 10; Conservative 2; Mismatches 3; Indels 23; Gaps 1;

QY 1 DEINQM-----HIDEXNE 15  
Db 81 DDIMNMTWQWKEISNYTGIIYLIIESQIQOEKNE 118

RESULT 12  
US-08-965-056-87  
Sequence 87, Application US/08965056  
Patent No. 6271198  
GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted  
APPLICANT: J. Kevin Judice  
APPLICANT: Robert S. McDowell  
APPLICANT: J. Christopher Phelan  
APPLICANT: Melissa A. Starovastnik  
APPLICANT: James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
TITLE OF INVENTION: Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/965,056  
FILING DATE: 05-No. 6271198-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-965-056-87

Query Match 45.9%; Score 39.5; DB 3; Length 198;  
Best Local Similarity 26.3%; Pred. No. 75;  
Matches 10; Conservative 2; Mismatches 3; Indels 23; Gaps 1;

QY 1 DEINQM-----HIDEXNE 15  
Db 150 DDIMNMTWQWKEISNYTGIIYLIIESQIQOEKNE 187

RESULT 13  
US-09-257-490-14  
Sequence 14, Application US/09257490A

Patent No. 6248328  
GENERAL INFORMATION:  
APPLICANT: Dietrich, Ursula  
APPLICANT: Von Briesen, Hagen  
APPLICANT: Grez, Manuel  
APPLICANT: Rudsamen-Waigmann, Helga  
TITLE OF INVENTION: HIV-1 virus isolates of a subtype and its differential  
TITLE OF INVENTION: diagnostics, a vaccine against HIV-1 virus infections  
TITLE OF INVENTION: of this subtype and method of producing same, use of  
FILE REFERENCE: 10496/P58512US1  
CURRENT APPLICATION NUMBER: US/09/257,490A  
CURRENT FILING DATE: 1999-02-25  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 613  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
FEATURE:  
OTHER INFORMATION: HIV-1 (Mal)  
US-09-257-490-14

Query Match 45.3%; Score 39.5; DB 3; Length 613;  
 Best Local Similarity 26.3%; Pred. No. 2.6e+02;  
 Matches 10; Conservative 2; Mismatches 3; Indels 23; Gaps 1;

Qy 1 DE1WNOH-----HIDEKNE 15  
 Db 514 DD1WNNMTWQWEKEISNYGTGIYLIRESQIQEKE 551

RESULT 14  
 US-09-621-976-4783  
 ; Sequence 4783, Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jobert, S.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 4783  
 ; LENGTH: 148  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-621-976-4783

Query Match 45.3%; Score 39; DB 4; Length 148;  
 Best Local Similarity 41.7%; Pred. No. 65;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 4 WNOHIDEKNE 15  
 Db 97 WOIHHTEDKDK 108

RESULT 15  
 US-09-533-029-88  
 ; Sequence 88, Application US/09533029  
 ; Patent No. 6664446  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Heard, Jacqueline  
 ; APPLICANT: Broun, Pierre  
 ; APPLICANT: Riechmann, Jose-Luis  
 ; APPLICANT: Keddie, James  
 ; APPLICANT: Pineda, Omaira  
 ; APPLICANT: Adam, Luc  
 ; APPLICANT: Samaha, Raymond  
 ; APPLICANT: Zhang, James  
 ; APPLICANT: Yu, Guo-Liang  
 ; APPLICANT: Ratcliffe, Oliver  
 ; APPLICANT: Pilgrim, Marsha  
 ; APPLICANT: Jiang, Cai-Zhong  
 ; APPLICANT: Reuber, Lynne  
 ; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES  
 ; FILE REFERENCE: MFI-010  
 ; CURRENT APPLICATION NUMBER: US/09/533,029  
 ; CURRENT FILING DATE: 2000-03-22  
 ; EARLIER APPLICATION NUMBER: 60/125,814  
 ; EARLIER FILING DATE: 1999-03-23  
 ; NUMBER OF SEQ ID NOS: 121  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 88  
 ; LENGTH: 226  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; OTHER INFORMATION: G221  
 US-09-533-029-88

Query Match 45.3%; Score 39; DB 4; Length 226;  
 Best Local Similarity 55.6%; Pred. No. 1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DE1WNOHMT 9  
 Db 214 DD1WSSKMTL 222

Search completed: July 29, 2004, 11:38:39  
 Job time : 14 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:30:55 ; Search time 31 Seconds

(without alignments)  
152.670 Million cell updates/sec

Title: US-10-040-128-1

Sequence: 1 DEIMNQMHIDEKNE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP\_invertebrate: \*  
6: SP\_mammal: \*  
7: SP\_mhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP\_unclassified: \*  
15: SP\_virus: \*  
16: SP\_bacteriap: \*  
17: SP\_archaeap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	73	84.9	782	6 Q8WNW5	Q8WNW5 sus scrofa
2	50	58.1	77	16 Q8Y643	Q8Y643 listeria mo
3	50	58.1	773	13 Q8AYD0	Q8AYD0 gallus gall
4	49	57.0	77	16 Q92AF1	Q92AF1 listeria in
5	47	54.7	712	5 Q20323	Q20323 caenorhabdi
6	46	53.5	241	11 Q9D996	Q9D996 mus musculu
7	46	53.5	520	8 Q8B916	Q8B916 pleistodont
8	46	53.5	933	4 Q8N768	Q8N768 homo sapien
9	46	53.5	1029	4 Q9C099	Q9C099 homo sapien
10	46	53.5	9341	5 Q813N9	Q813N9 plasmodium
11	45	52.3	176	10 Q8RYF1	Q8RYF1 arena strig
12	45	52.3	264	16 Q8YJ63	Q8YJ63 bradyrhizob
13	45	52.3	391	16 Q31773	Q31773 bacillus su
14	45	52.3	502	5 Q61644	Q61644 caenorhabdi
15	45	52.3	787	16 Q8F093	Q8F093 leptospira
16	45	52.3	847	11 Q9CUB1	Q9CUB1 mus musculu

17	45	52.3	942	16 Q87S90	Q87S90 vibrrio para
18	45	52.3	993	11 Q9DC06	Q9DC06 mus musculu
19	45	52.3	1682	4 Q15054	Q15054 homo sapien
20	45	52.3	2273	3 Q42823	Q42823 saccharomyc
21	44	51.2	558	5 Q20375	Q20375 caenorhabdi
22	44	51.2	586	5 Q95277	Q95277 caenorhabdi
23	43.5	50.6	1422	5 Q81S80	Q81S80 plasmodium
24	43	50.0	92	10 Q8LSF0	Q8LSF0 gossypium h
25	43	50.0	171	10 Q04666	Q04666 arabidopsis
26	43	50.0	243	6 Q7YR88	Q7YR88 sus scrofa
27	43	50.0	367	16 Q97KM0	Q97KM0 clostridium
28	43	50.0	425	16 Q7V243	Q7V243 prochloroco
29	43	50.0	444	16 Q99XC5	Q99XC5 staphylococ
30	43	50.0	487	16 Q83HS6	Q83HS6 tropheryma
31	43	50.0	487	16 Q83GF4	Q83GF4 tropheryma
32	43	50.0	712	5 Q81TC7	Q81TC7 plasmodium
33	43	50.0	876	15 Q9QS84	Q9QS84 human immun
34	43	50.0	900	10 Q9FM50	Q9FM50 arabidopsis
35	43	50.0	1710	5 Q81239	Q81239 plasmodium
36	43	50.0	2285	5 Q81240	Q81240 plasmodium
37	42.5	49.4	844	15 Q41789	Q41789 human immun
38	42.5	49.4	845	15 Q9GRB1	Q9GRB1 human immun
39	42.5	49.4	856	15 Q73296	Q73296 human immun
40	42	48.8	157	5 Q9NFS8	Q9NFS8 plasmodium
41	42	48.8	157	17 Q8PZJ1	Q8PZJ1 methanosarc
42	42	48.8	166	16 Q98F49	Q98F49 rhizobium l
43	42	48.8	237	16 Q7VFM5	Q7VFM5 helicobacte
44	42	48.8	296	16 Q7YR85	Q7YR85 synechococ
45	42	48.8	342	11 Q8C915	Q8C915 mus musculu

## ALIGNMENTS

## RESULT 1

Q8WNW5 PRELIMINARY; PRT; 782 AA.  
AC Q8WNW5;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE VB cadherin.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hirano M., Hirano K., Nishimura J., Kanaide H.;  
RT "Transcriptional up-regulation of p27Kip1 during contact-induced  
RT growth arrest in the endothelial cells.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DDJJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=99119339; PubMed=9918666;  
RA Hirano M., Nitro N., Hirano K., Nishimura J., Harshorne D.J.,  
RA Kanaide H.;  
RT "Expression, subcellular localization and cloning of the 130 kDa  
RT regulatory subunit of myosin phosphatase in porcine aortic endothelial  
RL cells.";  
RL Biochem. Biophys. Res. Commun. 254:490-496(1999).  
CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
CC SORTING OF HETEROGENEOUS CELL TYPES (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC EMBL; AB046120; BAB82983.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005194; F:cell adhesion molecule activity; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.



```

Db          62 DELMEAYLDEKKN 76

RESULT 5
Q20323      PRELIMINARY;      PRT;      712 AA.
AC Q20323;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN F42C5.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;

[1]
SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z.;
RT "The sequence of C. elegans cosmid F42C5.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40799; AAA81480.1; -.
DR PIR; T16338; T16338.
DR WormPep; F42C5.4; CE04557.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KM Hypothetical protein.
SQ SEQUENCE 712 AA; 84368 MW; 06A97BE962739FF6 CRC64;

Query Match          54.7%; Score 47; DB 5; Length 712;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY          1 DEINQMOMH 8
DB          341 DEINQMOMH 348

RESULT 6
Q9D996      PRELIMINARY;      PRT;      241 AA.
AC Q9D996;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 1700122011Rik protein.
GN 1700122011Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

[1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

```

```

RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barth G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayaehizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RX EMBL; AK007241; BAB24910.1; -.
DR MGD; MGI:1923901; 1700122011Rik.
SQ SEQUENCE 241 AA; 27288 MW; D2A47C328823597E CRC64;

Query Match          53.5%; Score 46; DB 11; Length 241;
Best Local Similarity 40.0%; Pred. No. 18;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY          1 DEINQMOMHIDEKNE 15
DB          38 EDINQMOMHIDEKNE 52

RESULT 7
Q9B916      PRELIMINARY;      PRT;      520 AA.
AC Q9B916;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Cytochrome oxidase subunit I (BC 1.9.3.1) (COI) (Cytochrome c oxidase
polypeptide I) (Fragment).
DE Pleistodontes rigisamos.
OS Pleistodontes rigisamos.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Chalcidoidea;
OC Agaonidae; Agaoninae; Pleistodontes.
OX NCBI_TaxID=108973;

[1]
SEQUENCE FROM N.A.
RC STRAIN=GW943;
RA Weiblen G.D.;
RT "Phylogenetic analyses of dioecious fig pollinators based on
mitochondrial DNA sequences and morphology.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DE FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
AND COPPER B (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
C + 2 H(2)O.
CC -1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
DR EMBL; AF200409; AAK00121.1; -.
DR HSSP; P18401; 1FFT.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005759; C:mitochondrion; IEA.

```

DR GO: GO:0009481; F:a3-lyc-type cytochrome c oxidase; IEA.  
 DR GO: GO:0009482; F:a3-lyc-type cytochrome c oxidase; IEA.  
 DR GO: GO:0009483; F:ca3-lyc-type cytochrome c oxidase; IEA.  
 DR GO: GO:0009484; F:cb3-lyc-type cytochrome c oxidase; IEA.  
 DR GO: GO:0009485; F:cb3-lyc-type cytochrome c oxidase; IEA.  
 DR GO: GO:0016429; F:cytochrome-c oxidase activity; IEA.  
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO: GO:0006118; P:electron transport; IEA.  
 DR GO: GO:0006110; P:transport; IEA.  
 DR InterPro: IPR00883; COX1.  
 DR Pfam: PF00115; COX1.1.  
 DR PRINTS: PR01165; CYCOXIDASE1.  
 DR PROSITE: PS00077; COX1.1.  
 DR Copper; Electron transport; Heme; Inner membrane; Membrane;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport;  
 KW Mitochondrion.  
 FT NON TER 1  
 SQ SEQUENCE 520 AA; 58939 MW; 315025C557C5A7F1 CRC64;

Query Match 53.5%; Score 46; DB 8; Length 520;  
 Best Local Similarity 52.9%; Pred. No. 41;  
 Matches 9; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 1 DEINWQMHIDEKNE 15  
 :|||:|:|:|:|:  
 Db 492 EIMWSPRIYIDEVND 508

RESULT 8  
 Q8N768 PRELIMINARY; PRT; 933 AA.  
 ID Q8N768  
 AC Q8N768;  
 DT 01-OCT-2002 (TREMblrel. 22, Created)  
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Hypothetical protein KIAA1764.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 GN NCB1\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Testis;  
 RA Strausberg R.;  
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC030701; AAH30701.1; -.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003885; LRR\_cyst.  
 DR InterPro: IPR007092; LRR\_SDS22.  
 DR Pfam: PF00560; LRR; 4.  
 DR SMART; SM00365; LRR\_SD22; 3.  
 KW Hypothetical protein.  
 SQ SEQUENCE 933 AA; 108165 MW; 66831C7105DAEB3C CRC64;

Query Match 53.5%; Score 46; DB 4; Length 933;  
 Best Local Similarity 46.7%; Pred. No. 76;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DEINWQMHIDEKNE 15  
 :|||:|:|:|:|:  
 Db 62 DHINWQHLDLSNQ 76

RESULT 9  
 Q9C099 PRELIMINARY; PRT; 1029 AA.  
 ID Q9C099  
 AC Q9C099;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
 DE Hypothetical protein KIAA1764 (Fragment).  
 GN KIAA1764.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCB1\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082932; PubMed=11214970;  
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,  
 RT "Prediction of the coding sequences of unidentified human genes. XIX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:347-355(2000).  
 DR EMBL; AB051551; BAB21855.1; -.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003603; LRRcap.  
 DR InterPro: IPR007092; LRR\_SDS22.  
 DR InterPro: IPR002017; Spectrin.  
 DR Pfam; PF00560; LRR; 4.  
 DR SMART; SM00446; LRRcap; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 1029 AA; 119884 MW; 1817DD4D446462DB CRC64;

Query Match 53.5%; Score 46; DB 4; Length 1029;  
 Best Local Similarity 46.7%; Pred. No. 85;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DEINWQMHIDEKNE 15  
 :|||:|:|:|:|:  
 Db 59 DHINWQHLDLSNQ 73

RESULT 10  
 Q8I3N9 PRELIMINARY; PRT; 9341 AA.  
 ID Q8I3N9  
 AC Q8I3N9;  
 DT 01-MAR-2003 (TREMblrel. 23, Created)  
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN PFE1120W.  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OC NCB1\_TaxID=36329;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,  
 RA Hall N., Bowman S., Churcher C., Quail M., Barrett B.;  
 RL Submitted (Sep-2002) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22255708; PubMed=12368867;  
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,  
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,  
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,  
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,  
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,  
 RA Fretwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,  
 RA Harper D., Hauser H., Hornsby T., Holyoak S., Horrocks P.,  
 RA Humphray S., Jagels K., James K.D., Johnson D., Kexhorou A.,  
 RA Knights A., Konfortov B., Kyles S., Mooney P., Moule S., Murphy L.,  
 RA line A., Maddison M., McLean J., Mooney P., Moyle S., Simmonds M.,  
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitch E.,  
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders S., Stevens K.,  
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
 RA Nulton J.E., Craig A., Newbold C., Barrett B.G.;  
 RT "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13.";  
 RL Nature 419:527-531(2002).  
 DR EMBL; AL929353; CAD51588.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 9341 AA; 1121395 MW; BAE88CDEBF66B92D CRC64;

Query Match 53.5%; Score 46; DB 5; Length 9341;  
 Best Local Similarity 42.9%; Pred. No. 8.6e+02;

Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 DEINQWIDEKN 14  
6064 NDIMNDIHINKTN 6077

RESULT 11  
Q8RYF1 PRELIMINARY; PRT; 176 AA.

AC Q8RYF1  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative resistance protein (Fragment).  
GN RGA.  
OS *Avena striigosa* (black oat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Avenae; Avena.  
OC NCBI\_TaxID=38783;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lorange Y., Irigoyen M.L., Fominaya A., Ferrer E.;  
RT "Resistance gene analogs in *Avena striigosa*.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ437574; CAD26862.1; -;  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0006915; P:apoptosis; IEA.  
DR InterPro; IPR002182; NB-ARC.  
DR Pfam; PF00931; NB-ARC; 1.  
FT NON\_TER 176 176  
FT SEQUENCE 176 AA; 20133 MW; D60982A15C6FE2F6 CRC64;  
SQ

Query Match 52.3%; Score 45; DB 10; Length 176;  
Best Local Similarity 58.3%; Pred. No. 19;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 INQWIDEKN 14  
62 LKMSNIDEKD 73

RESULT 12  
Q89JB9 PRELIMINARY; PRT; 264 AA.

AC Q89JB9  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE B115364 protein.  
GN B115364  
OS *Bradyrhizobium japonicum*.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OC NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=2248498; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uehimi T.,  
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
RT *Bradyrhizobium japonicum* USDA110.";  
RL DNA Res. 9:189-197 (2002).  
DR EMBL; AP005954; BAC50629.1; -;  
DR GO; GO:0005874; C:structural molecule; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0007018; P:microtubule-based movement; IEA.  
DR InterPro; IPR002453; Beta\_tubulin.

DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; 1.  
QW Complete proteome.  
SQ SEQUENCE 264 AA; 28370 MW; DFF604CED986052 CRC64;

Query Match 52.3%; Score 45; DB 16; Length 264;  
Best Local Similarity 42.9%; Pred. No. 30;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DEINQWIDEKN 14  
212 DQYMAAHVDEDMN 225

RESULT 13  
O31773 PRELIMINARY; PRT; 391 AA.

AC O31773  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Penicillin-binding protein.  
GN PBX.  
OS *Bacillus subtilis*.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OC NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=9804403; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertolo M.G., Bessieres P., Bolojin A., Borchert S.,  
RA Boriss R., Bourisler L., Brans A., Braun M., Brignall S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell I.B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entian K.D., Erttington J., Fabret C., Ferrati E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grand G.,  
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert B., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris H., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,  
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Prescan E., Pujic P., Purnelle R., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,  
RA Sekiguchi J., Sekowska A., Serron S.J., Serron P., Shin B.S.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpeira P., Tognoni A.,  
RA Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,  
RA Viari A., Wandut R., Wedler H., Wedler H., Weltzeneger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoko K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*.";  
RL Nature 390:249-256 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z99112; CAB1368.1; -;  
DR PIR; C69673; C69673.  
DR InterPro; IPR001466; Beta\_lactamase.  
DR Pfam; PF00144; beta-lactamase; 1.  
DR Complete proteome.  
SQ SEQUENCE 391 AA; 43858 MW; 08D3D7AB99826CCE CRC64;

Query Match 52.3%; Score 45; DB 16; Length 391;

Best Local Similarity 53.8%; Pred. No. 45;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 3 EIMNQHIDEKNE 15  
|||:|:|:  
Db 36 IYNAALHNSENE 48

## RESULT 14

061644 PRELIMINARY; PRT; 502 AA.  
ID 061644  
AC 061644  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Putative PP2C protein phosphatase FEM-2.  
GN FEM-2.  
OS Caenorhabditis briggsae.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6238;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AF16;  
RA Hansen D.; Pilgrim D.;  
RT "Molecular Evolution of a Sex Determination Protein: FEM-2 (PP2C) in  
RT Caenorhabditis."  
RL Genetics 0:0-0(1998).  
CC EMBL: AF054982; AAC08602.1; -  
DR GO: GO:0008287; C:protein serine/threonine phosphatase complex; IEA.  
DR GO: GO:0016787; F:hydrolase activity; IEA.  
DR GO: GO:0000287; F:magnesium ion binding; IEA.  
DR GO: GO:0004722; F:protein serine/threonine phosphatase activity; IEA.  
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.  
DR InterPro: IPR000222; PP2C.  
DR InterPro: IPR001932; PP2C-like.  
DR Pfam: PF00481; PP2C; 1.  
DR SMART: SM00332; PP2CC; 1.  
DR SMART: SM00331; PP2C-STG; 1.  
DR PROSITE: PS01032; PP2C; 1.  
KW Hydrolase; Magnesium.  
SQ SEQUENCE 502 AA; 57327 MW; 9EC5F1F8FF799A41 CRC64;  
Query Match 52.3%; Score 45; DB 5; Length 502;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 2 EIMNQHIDEKNE 15  
|||:|:|:  
Db 481 DLMRMKIDESDE 494

## RESULT 15

08F093 PRELIMINARY; PRT; 787 AA.  
ID 08F093  
AC 08F093  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Sulfatase family protein.  
GN IAS603.  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
OX NCBI\_TaxID=173;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
RA Ren S.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE01515; AAN50801.1; -  
DR GO: GO:0008484; F:sulfuric ester hydrolase activity; IEA.  
DR GO: GO:0008152; P:metabolism; IEA.

DR InterPro: IPR000917; Sulfatase.  
DR Pfam: PF00884; Sulfatase; 1.  
KW Complete proteome.  
SQ SEQUENCE 787 AA; 91636 MW; 1BF953631B28F798 CRC64;

Query Match 52.3%; Score 45; DB 16; Length 787;  
Best Local Similarity 61.5%; Pred. No. 93;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EIMNQHIDEKNE 14  
|||:|:|:  
Db 176 EIMNQHIDEKNE 188

Search completed: July 29, 2004, 11:37:08  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:30:30 ; Search time 6.66667 Seconds

(without alignments)  
117.158 Million cell updates/sec

Title: US-10-040-128-1  
Perfect score: 86  
Sequence: 1 DEIMNQWHDDEKNE 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	90.7	784	1	CAD5_MOUSE
2	73	84.9	782	1	CAD5_PIG
3	73	84.9	784	1	CAD5_HUMAN
4	45	52.3	418	1	GLYA_STRP3
5	45	52.3	418	1	GLYA_STRP8
6	45	52.3	418	1	GLYA_STRP1
7	45	52.3	418	1	GLYA_STRP1
8	43	50.0	435	1	HFA1_YEAST
9	42	48.8	363	1	HISX_BUCAR
10	42	48.8	363	1	DP3X_BUCAR
11	41	47.7	772	1	PHG1_HUMAN
12	40	46.5	146	1	CAD7_HUMAN
13	40	46.5	271	1	HBBC_HOPLI
14	40	46.5	358	1	PTPA_HUMAN
15	40	46.5	560	1	CALX_SCHPO
16	40	46.5	1113	1	HDA5_MOUSE
17	40	46.5	1122	1	HDA5_MOUSE
18	40	46.5	1812	1	SUTS_SUNMU
19	39.5	45.9	859	1	ENV_HV1MA
20	39.5	45.9	860	1	ENV_HV1MA
21	39	45.3	146	1	YERX_SCHPO
22	39	45.3	186	1	HBR_BRATR
23	39	45.3	395	1	SCM1_YEAST
24	39	45.3	401	1	BAF3_YEAST
25	39	45.3	603	1	VE1_HPV21
26	39	45.3	605	1	VE1_HPV21
27	39	45.3	621	1	SRPR_YEAST
28	39	45.3	684	1	HTPG_PORGI
29	39	45.3	753	1	TDR6_HUMAN
30	39	45.3	848	1	CYRA_YERIN
31	39	45.3	2136	1	YCP2_MARPO
32	38.5	44.8	863	1	ENV_HV128
33	38	44.2	135	1	RS9_PYRAB

## ALIGNMENTS

RESULT 1	CAD5_MOUSE	STANDARD;	PRT;	784 AA.
AC	P55284; 035542;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Vascular endothelial-cadherin precursor (VE-cadherin) (CDH5).			
GN	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain capillary;			
RX	MEDLINE=96141083; PubMed=8555485;			
RA	Breier G., Breivario F., Cavada L., Berthier R., Schnuerch H.,			
RT	Gotsch U., Westweder D., Rissau W., Dejama E.;			
RT	"Molecular cloning and expression of murine vascular endothelial-			
RT	cadherin in early stage development of cardiovascular system.";			
RL	Blood 87:630-641(1996).			
RN	[2]			
RP	REVIEWS TO 67-70.			
RA	Breviario F.;			
RL	Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A., AND FUNCTION.			
RC	STRAIN=BA1B/C; TISSUE=Breast carcinoma;			
RX	MEDLINE=97364256; PubMed=9220534;			
RA	Matsuyoshi N., Imamura S.;			
RT	Takeuchi M., Imamura S.;			
RT	"In vivo evidence of the critical role of cadherin-5 in murine			
RT	vascular integrity.";			
RT	Proc. Assoc. Am. Physicians 109:362-371(1997).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6; TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Klausner R.L., Feltingold E.A., Grouse L.H., Derge J.G.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,			
RA	Diachenko L., Martins K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bernaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carinici P., Prange C.,			
RA	Rah S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Boesk S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Falley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Mairra M.A.;			

34	38	44.2	135	1	RS9_PYRPU	Q8U0E7 pyrococcus
35	38	44.2	135	1	RS9_PYRPO	O59299 pyrococcus
36	38	44.2	146	1	HBE_HYLSY	Q95190 hylobates s
37	38	44.2	254	1	FTSQ_HAEIN	P45067 haemophilus
38	38	44.2	276	1	KDVI_BACHD	Q9XAX0 bacillus ha
39	38	44.2	311	1	Y467_MYCCE	P47705 mycoplasma
40	38	44.2	315	1	SPDE_NICSY	O48660 nicotiana s
41	38	44.2	315	1	SPDL_HYONI	O48660 nicotiana s
42	38	44.2	317	1	SPD2_DARTS	O48660 nicotiana s
43	38	44.2	342	1	SPDE_LYCES	Q96557 datura str
44	38	44.2	420	1	GLYA_STRMU	Q9Z45 lycopersico
45	38	44.2	427	1	CISY_SALTY	Q8U0E7 streptococc
						O68883 salmonella

RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X83930; CAAS8782.2; -;  
 CC EMBL; D63942; BAA2617.1; -;  
 CC EMBL; BC054790; AAH54790.1; -;  
 CC HSSP; P15116; INCU.  
 CC MGD; MGI:105057; Cdh5.  
 CC InterPro; IPR002126; Cadherin.  
 CC InterPro; IPR000233; Cadherin\_C\_term.  
 CC Pfam; PF00028; cadherin\_5.  
 CC Pfam; PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS; PR00205; CADHERIN.  
 CC SMART; SM00112; CA; 5.  
 CC PROSITE; PS00232; CADHERIN\_1; 3.  
 CC PROSITE; PS50268; CADHERIN\_2; 5.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT PROPEP 1 24 POTENTIAL.  
 FT SIGNAL 25 45 POTENTIAL.  
 FT CHAIN 46 784 VASCULAR ENDOTHELIAL-CADHERIN.  
 FT TRANSMEM 46 593 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 594 620 POTENTIAL.  
 FT DOMAIN 621 784 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 46 149 CADHERIN 1.  
 FT DOMAIN 150 256 CADHERIN 2.  
 FT DOMAIN 257 371 CADHERIN 3.  
 FT DOMAIN 372 476 CADHERIN 4.  
 FT DOMAIN 477 593 CADHERIN 5.  
 FT DOMAIN 738 753 SER-RICH.  
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 784 AA; 87902 MW; 7B75698DE2F7E160 CRC64;  
 Query Match Score 78; DB 1; Length 784;  
 Best Local Similarity 93.3%; Pred. No. 5, 9e-05;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DEINQMHIIDEKN 15  
 DB 46 DWINQMHIIDEKN 60  
 RESULT 2  
 CAD5\_PIG STANDARD; PRT; 782 AA.  
 AC 002840;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5).  
 GN CDH5.  
 OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OC NCBI\_TaxID=9823;  
 RN (1)  
 RP Kishaw P.J.;  
 RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions. It  
 CC associates with alpha-catenin forming a link to the cytoskeleton  
 CC (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Found at cell-cell  
 CC boundaries and probably at cell-matrix boundaries (by similarity).  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Y13919; CAA7425.1; -;  
 CC HSSP; P09803; IEDH.  
 CC InterPro; IPR002126; Cadherin.  
 CC InterPro; IPR000233; Cadherin\_C\_term.  
 CC Pfam; PF00028; cadherin\_5.  
 CC Pfam; PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS; PR00205; CADHERIN.  
 CC SMART; SM00112; CA; 5.  
 CC PROSITE; PS00232; CADHERIN\_1; 3.  
 CC PROSITE; PS50268; CADHERIN\_2; 5.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT PROPEP 1 22 POTENTIAL.  
 FT SIGNAL 23 44 POTENTIAL.  
 FT CHAIN 45 782 VASCULAR ENDOTHELIAL-CADHERIN.  
 FT TRANSMEM 45 592 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 593 619 POTENTIAL.  
 FT DOMAIN 620 782 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 45 148 CADHERIN 1.  
 FT DOMAIN 149 255 CADHERIN 2.  
 FT DOMAIN 256 370 CADHERIN 3.  
 FT DOMAIN 371 475 CADHERIN 4.  
 FT DOMAIN 476 592 CADHERIN 5.  
 FT DOMAIN 736 751 SER-RICH.  
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 782 AA; 87546 MW; 7403F974E2DF782F CRC64;  
 Query Match Score 73; DB 1; Length 782;  
 Best Local Similarity 92.9%; Pred. No. 0.00038;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DEINQMHIIDEKN 14  
 DB 45 DWINQMHIIDEKN 58  
 RESULT 3  
 CAD5\_HUMAN STANDARD; PRT; 784 AA.  
 AC P3151;  
 DT 01-OCT-1993 (Rel. 27, Created)



DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DB Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5)  
 DE (7B4 antigen) (CD144 antigen).  
 GN CDH5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Endothelial cells;  
 RX MEDLINE=95353875; PubMed=7627717;  
 RA Breviaro F., Cavada L., Corada M., Martin-Padura I., Navarro P.,  
 RA Goley J., Introna M., Gulino D., Lampugnani M.G., Dejana E.;  
 RT "Functional properties of human vascular endothelial cadherin  
 RT (7B4/cadherin-5), an endothelium-specific cadherin,"  
 RL Arterioscler. Thromb. Biol. 15:1229-1239(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=97362755; PubMed=9219219;  
 RA Ali U., Liao F., Martens E., Muller W.A.;  
 RT "Vascular endothelial cadherin (VE-cadherin): cloning and role in  
 RT Microcirculation cell-cell adhesion,"  
 RL Microcirculation 4:267-277(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=10861224;  
 RA Shintoyama Y., Tsujimoto G., Kitajima M., Natori M.;  
 RT "Identification of three human type-II classic cadherins and frequent  
 RT heterophilic interactions between different subclasses of type-II  
 RT classic cadherins,"  
 RL Biochem. J. 349:159-167(2000).  
 RN [4]  
 RP SEQUENCE OF 5-784 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=91283540; PubMed=2059658;  
 RA Suzuki S., Sano K., Tanihara H.;  
 RT "Diversity of the cadherin family: evidence for eight new cadherins  
 RT in nervous tissue,"  
 RL Cell Regul. 2:261-270(1991).  
 RN [5]  
 RP PARTIAL SEQUENCE.  
 RC TISSUE=Endothelial cells;  
 RX MEDLINE=92394977; PubMed=1522121;  
 RA Lampugnani M.G., Resnati M., Raiteri M., Pigott R., Pisacane A.,  
 RA Hohen G., Ruco L.P., Dejana E.;  
 RT "A novel endothelial-specific membrane protein is a marker of  
 RT cell-cell contacts,"  
 RL J. Cell Biol. 118:1511-1522(1992).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions. It  
 CC associates with alpha-catenin forming a link to the cytoskeleton.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Found at cell-cell  
 CC boundaries and probably at cell-matrix boundaries.  
 CC -1- TISSUE SPECIFICITY: Endothelial tissues and brain.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -1- DATABASE: NAME=PROV. NOTE=CD guide CD144 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd144.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

DR EMBL: X79981; CAA56306.1; -  
 DR EMBL: U84722; AAB41796.1; -  
 DR EMBL: AB035304; BAA87418.1; -  
 DR EMBL: X59796; CAA42468.1; -  
 DR PIR: S49893; ITHUC5.  
 DR HSSP: P15116; INCU.  
 DR GlycoStidB: P33151; -  
 DR Genew: HGNC:1764; CDH5.  
 DR MIM: 601120; -  
 DR GO: GO:0005624; C:membrane fraction; TAS.  
 DR GO: GO:0005886; C:plasma membrane; TAS.  
 DR GO: GO:0007156; P:homophilic cell adhesion; TAS.  
 DR InterPro: IPR002126; Cadherin.  
 DR InterPro: IPR000233; Cadherin\_C\_term.  
 DR Pfam: PF01049; cadherin; 5-term.  
 DR PRINTS: PRO0205; CADHERIN.  
 DR SMART: SM00112; CA; 5.  
 DR PROSITE: PS00232; CADHERIN\_1; 3.  
 DR PROSITE: PS0268; CADHERIN\_2; 5.  
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT SIGNAL. 1 25  
 FT PROPEP. 26 47  
 FT CHAIN 48 784  
 FT DOMAIN 48 593  
 FT TRANSMEM 594 620  
 FT DOMAIN 621 784  
 FT DOMAIN 48 151  
 FT DOMAIN 152 258  
 FT DOMAIN 259 372  
 FT DOMAIN 373 477  
 FT DOMAIN 478 593  
 FT DOMAIN 736 753  
 FT CARBOHYD 61 61  
 FT CARBOHYD 112 112  
 FT CARBOHYD 157 157  
 FT CARBOHYD 362 362  
 FT CARBOHYD 442 442  
 FT CARBOHYD 523 523  
 FT CARBOHYD 535 535  
 FT CONFLICT 517 517  
 SQ SEQUENCE 784 AA; 87516 MW; F643BFC222A599DE CRC64;  
 Query Match 84.9%; Score 73; DB 1; Length 784;  
 Best Local Similarity 92.9%; Pred. No. 0.00038;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DEINQMHIDEKN 14  
 Db 48 DWINQMHIDEKN 61  
 ID GLYA\_STRP3 STANDARD; PRT; 418 AA.  
 AC Q8K7H8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serine hydroxymethyltransferase (EC 2.1.2.1) (Serine methylase)  
 DE (SMT).  
 GN GLYA OR SPYM3 0803 OR SPS1002.  
 OS Streptococcus pyogenes (serotype M3).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=198466;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MGA5315 / Serotype M3;  
 RX MEDLINE=22133808; PubMed=12122206;  
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,  
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,

RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,  
 RA Schlievert P.M., Musser J.M.;  
 RA "Genome sequence of a serotype M3 strain of group A Streptococcus:  
 RT phase-encoded toxins, the high-virulence phenotype, and clone  
 RT emergence";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=SSI-1 / Serotype M3;  
 RX MEDLINE=22683278; PubMed=12799345;  
 RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,  
 RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,  
 RA Hayashi H., Hattori M., Hamada S.,  
 RA "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a  
 RT large-scale genomic rearrangement in invasive strains and new insights  
 RT into phage evolution";  
 RT Genome Res. 13:1042-1055(2003).  
 CC -1- FUNCTION: Interconversion of serine and glycine.  
 CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + glycine +  
 CC H(2)O = tetrahydrofolate + L-serine.  
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
 CC -1- PATHWAY: Key enzyme in the biosynthesis of purines, lipids,  
 CC hormones and other components.  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the SHMT family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE014152; AAM79410.1; -;  
 DR EMBL; AP005144; BAC64097.1; ALT\_INIT.  
 DR HAMAP; MF\_00051; -; 1  
 DR InterPro; IPR001085; Gly\_HyMettransf.  
 DR Pfam; PF00464; SHMT; 1.  
 DR PROSITE; PS00096; SHMT; 1.  
 KW Transferase; Pyridoxal phosphate; One-carbon metabolism;  
 KW Complete proteome.  
 FT BINDING 230 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT SEQUENCE 418 AA; 45079 MW; 1C3B5E167C13650 CRC64;  
 SQ  
 Query Match 52.3%; Score 45; DB 1; Length 418;  
 Best Local Similarity 42.9%; Pred. No. 7;  
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Oy 2 EINOMHIDEKNE 15  
 |:::|:::|  
 Db 14 EIMDAIHAERROE 27

RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
 RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Portella S.F.,  
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;  
 RA "Genome sequence and comparative microarray analysis of serotype M18  
 RT group A Streptococcus strains associated with acute rheumatic fever  
 RT outbreaks";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=SF370 / ATCC 700294 / Serotype M1;  
 RX MEDLINE=21192684; PubMed=11296296;  
 RA Ferretti J.J., McShan W.M., Ajdic D.U., Savic D.J., Savic G., Lyon K.,  
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 CC -1- FUNCTION: Interconversion of serine and glycine.  
 CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + glycine +  
 CC H(2)O = tetrahydrofolate + L-serine.  
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE010036; AAL97727.1; -;  
 DR HAMAP; MF\_00051; -; 1  
 DR InterPro; IPR001085; Gly\_HyMettransf.  
 DR Pfam; PF00464; SHMT; 1.  
 DR PROSITE; PS00096; SHMT; 1.  
 KW Transferase; Pyridoxal phosphate; One-carbon metabolism;  
 KW Complete proteome.  
 FT BINDING 230 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT SEQUENCE 418 AA; 45065 MW; ASE8A48CD419CASA CRC64;  
 SQ  
 Query Match 52.3%; Score 45; DB 1; Length 418;  
 Best Local Similarity 42.9%; Pred. No. 7;  
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Oy 2 EINOMHIDEKNE 15  
 |:::|:::|  
 Db 14 EIMDAIHAERROE 27

```

CC -1- PATHWAY: Key enzyme in the biosynthesis of purines, lipids,
CC hormones and other components.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the SHMT family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006556; AAK34017.1; -.
CC DR HSSP; P00477; IDPO.
CC DR HAMAP; MF_00051; -.
CC DR InterPro; IPR001085; Gly_HyMettransf.
CC DR Pfam; PF00464; SHMT; 1.
CC DR PROSITE; PS00096; SHMT; 1.
CC DR Transferase; Pyridoxal phosphate; One-carbon metabolism;
CC Complete proteome.
CC FT BINDING 230 230 PYRIDOXAL PHOSPHATE (By similarity).
CC SQ SEQUENCE 418 AA; 45096 MW; AF475E79D41D9A1F CRC64;
CC -----
CC Query Match 52.3%; Score 45; DB 1; Length 418;
CC Best Local Similarity 42.9%; Pred. No. 7;
CC Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 2 DEIMNQHIDEKNE 15
CC |:::|:::|
CC Db 14 ELMDAIHAEEROE 27
CC -----
CC RESULT 7
CC HPA1_YEAST STANDARD; PRT; 2273 AA.
CC ID HPA1_YEAST
CC AC P32874;
CC DT 01-OCT-1993 (Rel. 27, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE HPA1 protein.
CC GN HPA1 OR YMR207C OR YMR261.01C OR YMR325.08C.
CC OS Saccharomyces cerevisiae (Baker's Yeast).
CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CC OX NCBI_TaxID=4932;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=S288C / AB972;
CC RX MEDLINE=97133268; PubMed=9169872;
CC RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
CC Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
CC Jagels K., Lye G., Moutle S., Odell C., Pearson D., Rajandream M.A.,
CC Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
CC RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
CC XIII."
CC RL Nature 387:90-93(1997).
CC RN [2]
CC RP SEQUENCE OF 125-949 FROM N.A.
CC RX MEDLINE=94146412; PubMed=7906156;
CC RA Kearsey S.E.;
CC RT "Identification of a Saccharomyces cerevisiae gene closely related to
CC PAF3 (acetyl-CoA carboxylase)."
CC RL DNA Seq. 4:69-70(1993).
CC CC -1- COFACTOR: Biotin (By similarity).
CC CC -1- SIMILARITY: STRONG, TO ACETYL-COA CARBOXYLASE.
CC CC -1- CAUTION: The reading frame from which this protein in translated
CC has no Met initiation codon near to the 5' end. It does not seem to
CC be a pseudogene. There are no apparent frameshifts.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z49809; CAA69922.1; -.
CC DR EMBL; Z48755; CAA68647.1; -.
CC DR EMBL; Z22558; CAA60280.1; -.
CC DR PIR; S55089; S55089.
CC DR HSSP; P24182; IDVI.
CC DR Germonline; 142893; -.
CC DR SGD; S0004820; HPA1.
CC DR InterPro; IPR001882; Biotin BS.
CC DR InterPro; IPR005482; Biotin carb C.
CC DR InterPro; IPR000089; Biotin_lipoyl.
CC DR InterPro; IPR000022; Carboxyl trans.
CC DR InterPro; IPR005479; Cpase_L_D2.
CC DR InterPro; IPR005481; Cpase_L_N.
CC DR Pfam; PF02785; Biotin carb C; 1.
CC DR Pfam; PF00364; biotin_lipoyl; 1.
CC DR Pfam; PF01039; Carboxyl trans; 1.
CC DR Pfam; PF00289; Cpase_L_chain; 1.
CC DR Pfam; PF02786; Cpase_L_D2; 1.
CC DR PROSITE; PS00188; BIOTIN; 1.
CC DR PROSITE; PS00867; CPASE_1; 1.
CC DR PROSITE; PS00866; CPASE_2; 1.
CC DR Biotin; Ligase; ATP-binding.
CC FT NP BIND 332 337 ATP (By similarity).
CC FT ACT SITE 459 459 By similarity.
CC FT BINDING 804 804 BIOTIN (By similarity).
CC FT CONFLICT 661 661 F -> L (IN REF. 2).
CC SQ SEQUENCE 2273 AA; 259160 MW; 08727A301549DA92 CRC64;
CC -----
CC Query Match 52.3%; Score 45; DB 1; Length 2273;
CC Best Local Similarity 53.8%; Pred. No. 43;
CC Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 1 DEIMNQHIDEK 13
CC |:::|:::|
CC Db 1248 DEVEQHIIPEER 1260
CC -----
CC RESULT 8
CC H1SX_BUCAL STANDARD; PRT; 435 AA.
CC ID H1SX_BUCAL
CC AC P57201;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Histidinol dehydrogenase (EC 1.1.1.23) (HHD).
CC GN H1SD OR BU100.
CC OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
CC symbiotic bacterium).
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Buchnera.
CC OX NCBI_TaxID=118099;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Tokyo 1998;
CC RX MEDLINE=20445173; PubMed=10993077;
CC RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
CC RT "Genome sequence of the endocellular bacterial symbiont of aphids
CC Buchnera sp. Aps."
CC RL Nature 407:81-86(2000).
CC CC -1- FUNCTION: Catalyzes the sequential NAD-dependent oxidations of L-
CC histidinol to L-histidinolaldehyde and then to L-histidine (By
CC similarity).
CC CC -1- CATALYTIC ACTIVITY: L-histidinol + 2 NAD(+) + H(2)O = L-histidine
CC + 2 NADH.
CC CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC CC -1- PATHWAY: Histidine biosynthesis; ninth (last) step.
CC -1- SUBUNIT: Homodimer (By similarity).

```

```

CC -1- SIMILARITY: Belongs to the histidinol dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB001118; BAB12819.1; -.
CC HAMAP: MF_01024; -.
CC InterPro: IPR001692; Histidinol_dh.
CC Pfam: PF00815; Histidinol_dh_1.
CC PRINTS: PR00083; HOLIDHGRASE.
CC ProDom: PD002680; Histidinol_dh; 1.
CC TIGRFAMs: TIGR00069; hsd; 1.
CC PROSITE: PS00611; HISOL_DEHYDROGENASE; 1.
CC Histidine biosynthesis; Oxidoreductase; NAD; Metal-binding; Zinc;
KW Complete proteome.
FT ACT_SITE 327 BY SIMILARITY.
FT ACT_SITE 328 BY SIMILARITY.
FT METAL 260 ZINC (BY SIMILARITY).
FT METAL 263 ZINC (BY SIMILARITY).
FT METAL 361 ZINC (BY SIMILARITY).
FT METAL 420 ZINC (BY SIMILARITY).
SQ SEQUENCE 435 AA; 47843 MW; 83DD3DC4E43430E CRC64;

Query Match 50.0%; Score 43; DB 1; Length 435;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WNNQNHIDEK 13
Db 10 WNKLPDDEOK 19

RESULT 9
DP3X_BUCAP STANDARD; PRT; 363 AA.
AC Q8K983;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase III subunit gamma (EC 2.7.7.7).
GN DNAX OR BUSG46.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klason L., Canback B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
RA "50 million years of genomic strais in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
CC -1- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII complex. PolIII associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB014123; AAM68009.1; -.
CC InterPro: IPR003593; AAA_Atpase.
CC InterPro: IPR003959; AAA_Atpase_cent.
CC InterPro: IPR001270; Chadrin_cipa/B.
CC InterPro: IPR008921; Pol_clamp_load_C.
CC InterPro: IPR000862; Pfdomain.
CC Pfam: PF00004; AAA_1.
CC PRINTS: PR00300; CLIPROTASER.
CC SMART: SM00382; AAA; 1.
CC Transferase; DNA-directed DNA polymerase; DNA replication;
KW ATP-binding; Complete proteome.
FT NP_BIND 45 ATP (POTENTIAL).
SQ SEQUENCE 363 AA; 42186 MW; F22762B7493B091A CRC64;

Query Match 48.8%; Score 42; DB 1; Length 363;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 EWNQNHIDEKNE 15
Db 300 KIMNOLFKNENE 313

RESULT 10
PHG1_HUMAN STANDARD; PRT; 1113 AA.
ID PHG1_HUMAN
AC Q9ULU1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pleckstrin homology domain containing family G member 1 (Fragment).
DE PLEKHG1 OR KIAA1203.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirogawa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC -1- SIMILARITY: Contains 1 PH domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB033035; BAA6523.1; -.
CC Genew; HGNC:20884; PLEKHG1.
CC InterPro: IPR001849; PH.
CC InterPro: IPR000219; RhoGEF.
CC Pfam: PF00169; PH; 1.
CC SMART: SM00233; PH; 1.
CC PROSITE: PS50003; PH_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 144
FT DOMAIN 1
SQ SEQUENCE 1113 AA; 124624 MW; 62E8E2C698B3519A CRC64;

Query Match 48.8%; Score 42; DB 1; Length 1113;

```

Best Local Similarity 53.3%; Pred. No. 61;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DEIMNQHIDEKKE 15  
Db 564 DEIMNDLENYIKKE 578

RESULT 11  
CADJ\_HUMAN STANDARD; PRT; 772 AA.

AC Q9H159;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Cadherin-19 precursor (UniQ478/PRO941).  
GN CDH19 OR CDH12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20453188; PubMed=10995570;  
RA Koels P., Van Imschoot G., Van Roy F.;  
RT "Characterization of three novel human cadherin genes (CDH7, CDH19,  
RT and CDH20) clustered on chromosome 18q22-q23 and with high homology  
RT to chicken cadherin-7";  
RL Genomics 66:283-295(2000).  
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
CC They preferentially interact with themselves in a homophilic  
CC manner in connecting cells; cadherins may thus contribute to the  
CC sorting of heterogeneous cell types.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed in many tissues, with the exception  
CC of uterus.  
CC -1- SIMILARITY: Contains 5 cadherin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AJ007607; CAC13126.1; -  
DR EMBL; AY358654; AAC89017.1; -  
DR HSSP; P15116; INCT.  
DR Genew; HGNC:1758; CDH19.  
DR MIM; 603016; -  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0007156; P:homophilic cell adhesion; NAS.  
DR InterPro; IPR002126; Cadherin.  
DR InterPro; IPR000233; Cadherin\_C\_term.  
DR Pfam; PF00028; cadherin; 5.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SM00112; CA; 5.  
DR PROSITE; PS00232; CADHERIN\_1; 2.  
DR PROSITE; PS50268; CADHERIN\_2; 5.  
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
FT SIGNAL.  
FT SIGNAL 1 21  
FT PROPEP 22 43  
FT CHAIN 44 772  
FT DOMAIN 44 596  
FT TRANSEM 597 617  
FT DOMAIN 618 772  
FT DOMAIN 44 148  
FT DOMAIN 149 256  
FT DOMAIN 257 370  
FT DOMAIN 371 470

FT DOMAIN 470 581  
FT CARBOHYD 57 57  
FT CARBOHYD 74 74  
FT CARBOHYD 419 419  
FT CARBOHYD 437 437  
FT CARBOHYD 508 508  
FT CARBOHYD 515 515  
FT CARBOHYD 516 516  
FT CARBOHYD 534 534  
SQ SEQUENCE 772 AA; 87002 MW; 650AD27480343C39 CRC64;  
CADHERIN 5.  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 47.7%; Score 41; DB 1; Length 772;  
Best Local Similarity 50.0%; Pred. No. 60;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 INNQHIDEK 14  
Db 46 VNNQFVPEK 57

RESULT 12

ID HBBC\_HOPLI STANDARD; PRT; 146 AA.

AC P82316;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hemoglobin cathectic, beta chain (Hb(Ca) beta chain).  
GN HBB.  
OS Hoplosternum littorale (Haasari).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;  
OC Callichthyidae; Hoplosternum.  
OX NCBI\_TaxID=114109;  
RN [1]  
RP SEQUENCE, SUBUNIT, AND MASS SPECTROMETRY.  
RC TISSUE=Blood;  
RX MEDLINE=20298792; PubMed=10747999;  
RA Weber R.E., Fago A., Val A.L., Bang A., Van Hauwaert M.-L.,  
RA Deswilde S., Zal F., Moens L.;  
RT "Isohemoglobin differentiation in the bimodal-breathing amazon  
RT catfish Hoplosternum littorale";  
RL J. Biol. Chem. 275:17297-17305(2000).  
CC -1- FUNCTION: Involved in oxygen transport from the lung to the  
CC various peripheral tissues.  
CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.  
CC -1- TISSUE SPECIFICITY: Red blood cells.  
CC -1- MASS SPECTROMETRY: MW=15978; MW\_ERR=0.2; METHOD=Electrospray.  
CC -1- MISCELLANEOUS: This fish has two hemoglobins: cathectic and anodic.  
CC The cathectic Hb and anodic Hb display small and large Bohr effects  
CC respectively. In addition, the cathectic Hb displays a reverse Bohr  
CC effect and appreciable phosphate effects.  
CC -1- SIMILARITY: Belongs to the globin family.

HSSP; P02142; IOUT.  
DR InterPro; IPR002337; Beta.haem.  
DR InterPro; IPR000971; Globin.  
DR Pfam; PF00042; Globin; 1.  
DR PRINTS; PR00814; BETAHAEM.  
DR PROSITE; PS01033; GLOBIN; 1.  
KW Heme; Oxygen transport; Transport; Erythrocyte.  
FT METAL 63 63  
FT METAL 92 92  
SQ SEQUENCE 146 AA; 15976 MW; 4D75EB9FC8D3539 CRC64;  
IRON (HEME DISTAL LIGAND).  
IRON (HEME PROXIMAL LIGAND).

Query Match 46.5%; Score 40; DB 1; Length 146;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 INNQHIDE 11  
Db 14 INNQHIDE 22

RESULT 13  
 AROE\_CLOPE STANDARD; PRT; 271 AA.  
 ID AROE\_CLOPE STANDARD; PRT; 271 AA.  
 AC 08XMTB;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Shikimate 5-dehydrogenase (EC 1.1.1.25).  
 OS AROE OR CPE0700.  
 OC Clostridium perfringens.  
 OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13 / Type A;  
 RX MEDLINE=21664373; PubMed=11792842;  
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 RT flesh-eater.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
 CC -1- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 5-dehydroshikimate +  
 CC NADPH.  
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC fourth step.  
 CC -1- SIMILARITY: Belongs to the shikimate dehydrogenase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AP003187; BAB80406.1; -  
 CC HAMAP: MF\_00222; -; 1.  
 DR InterPro: IPR006152; Shikimate.  
 DR InterPro: IPR006151; Shikimate.DH.  
 DR Pfam: PF01488; Shikimate.DH; 1.  
 DR TIGRPFAMs: TIGR00507; aroE; 1.  
 KW Aromatic amino acid biosynthesis; Oxidoreductase; NADP;  
 KW Complete proteome.  
 SQ SEQUENCE 271 AA; 30505 MW; 311AF6C6B63CFE CRC64;  
 Query Match 46.5%; Score 40; DB 1; Length 271;  
 Best Local Similarity 40.0%; Pred. No. 28;  
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 DEINMOHIDEKNE 15  
 Db 247 EIRWNGIKVDKSTEE 261  
 RESULT 14  
 PTPA\_HUMAN STANDARD; PRT; 358 AA.  
 ID PTPA\_HUMAN STANDARD; PRT; 358 AA.  
 AC Q15257; Q15258; Q9BXK1; Q9NNZ7; Q9NNZ9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Protein phosphatase 2A, regulatory subunit B' (PP2A, subunit B', PR53  
 DE isoform) (phosphotyrosyl phosphatase activator) (PTPA).  
 GN PPP2R4 OR PTPA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (ISOFORM 1).  
 RC TISSUE=Heart;  
 RX MEDLINE=94253154; PubMed=8195217;

RA Cayla X., Van Hoof C., Bosch M., Waelkens E., Peeters B.,  
 RA Merlevede W., Goris J.;  
 RT "Molecular cloning, expression, and characterization of PTPA, a  
 RT protein that activates the tyrosyl phosphatase activity of protein  
 RT phosphatase 2A.";  
 RL J. Biol. Chem. 269:15668-15675(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Blood;  
 RX MEDLINE=96015056; PubMed=8530035;  
 RA Van Hoof C., Aly M., Garcia A., Cayla X., Cassiman J., Merlevede W.,  
 RA Goris J.;  
 RT "Structure and chromosomal localization of the human gene of the  
 RT phosphotyrosyl phosphatase activator (PTPA) of protein phosphatase  
 RT 2A.";  
 RL Genomics 28:261-272 (1995).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).  
 RX MEDLINE=20341117; PubMed=10880964;  
 RA Janssens V., van Hoof C., Martens E., de Baere I., Merlevede W.,  
 RA Goris J.;  
 RT "Identification and characterization of alternative splice products  
 RT encoded by the human phosphotyrosyl phosphatase activator gene.";  
 RL Eur. J. Biochem. 267:4406-4413(2000).  
 RL [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bluetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshylyki S., Carninci P., Muliyil S.J.,  
 RA Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Reversibly stimulates the variable phosphotyrosyl  
 CC phosphatase activity of PP2A core heterodimer in presence of ATP  
 CC and Mg(2+) (in vitro).  
 CC -1- SUBUNIT: Associates with PP2A heterodimeric core enzyme, composed  
 CC of a 36 kDa catalytic subunit (subunit C) and a 65 kDa constant  
 CC regulatory subunit (PR65 or subunit A).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=2; Synonyms=Beta;  
 CC IsoId=Q15257-1; Sequence=Displayed;  
 CC Name=1; Synonyms=Alpha;  
 CC IsoId=Q15257-2; Sequence=VSP\_005123;  
 CC Name=3; Synonyms=Delta;  
 CC IsoId=Q15257-3; Sequence=VSP\_005122;  
 CC Name=4; Synonyms=Epsilon;  
 CC IsoId=Q15257-4; Sequence=VSP\_005124;  
 CC -1- TISSUE SPECIFICITY: Widely expressed.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL; X73478; CAA51873.1; -
DR EMBL; X86428; CAA60163.1; -
DR EMBL; X86429; CAA60163.1; JOINED.
DR EMBL; X86430; CAA60163.1; JOINED.
DR EMBL; X86432; CAA60163.1; JOINED.
DR EMBL; X86434; CAA60163.1; JOINED.
DR EMBL; X86435; CAA60163.1; JOINED.
DR EMBL; X86436; CAA60163.1; JOINED.
DR EMBL; X86437; CAA60163.1; JOINED.
DR EMBL; X86438; CAA60163.1; JOINED.
DR EMBL; X86439; CAA60163.1; JOINED.
DR EMBL; X86428; CAB77601.1; -
DR EMBL; X86429; CAB77601.1; JOINED.
DR EMBL; X86430; CAB77601.1; JOINED.
DR EMBL; X86431; CAB77601.1; JOINED.
DR EMBL; X86432; CAB77601.1; JOINED.
DR EMBL; X86433; CAB77601.1; JOINED.
DR EMBL; X86434; CAB77601.1; JOINED.
DR EMBL; X86435; CAB77601.1; JOINED.
DR EMBL; X86436; CAB77601.1; JOINED.
DR EMBL; X86437; CAB77601.1; JOINED.
DR EMBL; X86438; CAB77601.1; JOINED.
DR EMBL; X86439; CAB77601.1; JOINED.
DR EMBL; X86428; CAB77602.1; -
DR EMBL; X86429; CAB77602.1; JOINED.
DR EMBL; X86430; CAB77602.1; JOINED.
DR EMBL; X86431; CAB77602.1; JOINED.
DR EMBL; X86432; CAB77602.1; JOINED.
DR EMBL; X86433; CAB77602.1; JOINED.
DR EMBL; X86434; CAB77602.1; JOINED.
DR EMBL; X86435; CAB77602.1; JOINED.
DR EMBL; X86436; CAB77602.1; JOINED.
DR EMBL; X86437; CAB77602.1; JOINED.
DR EMBL; X86438; CAB77602.1; JOINED.
DR EMBL; X86439; CAB77602.1; JOINED.
DR EMBL; X86428; CAB77603.1; -
DR EMBL; X86429; CAB77603.1; JOINED.
DR EMBL; X86430; CAB77603.1; JOINED.
DR EMBL; X86431; CAB77603.1; JOINED.
DR EMBL; X86432; CAB77603.1; JOINED.
DR EMBL; X86433; CAB77603.1; JOINED.
DR EMBL; X86434; CAB77603.1; JOINED.
DR EMBL; X86435; CAB77603.1; JOINED.
DR EMBL; X86436; CAB77603.1; JOINED.
DR EMBL; X86437; CAB77603.1; JOINED.
DR EMBL; X86438; CAB77603.1; JOINED.
DR EMBL; X86439; CAB77603.1; JOINED.
DR EMBL; BC002545; AAH02545.1; -
DR EMBL; BC011605; AAH1605.1; -
DR PIR; A54021; A54021.
DR Genew; HGNC:9308; PPP2R4.
DR MIM; 600756; -
DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0008601; F:protein phosphatase type 2A, intrinsic regu. ; NAS.
DR GO; GO:0008160; F:protein tyrosine phosphatase activator acti. ; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; NAS.
DR InterPro; IPR004327; Phc1tyr_phosase_ac.
DR Pfam; PF03095; PPPA; 1.
KW Alternative splicing.
FT VARSPLIC 45 108 Missing (in isoform 3).
FT VARSPLIC 73 107 Missing (in isoform 1).
FT VARSPLIC 73 149 Missing (in isoform 4).
FT VARSPLIC 73 149 Missing (in isoform 4).
FT CONFLICT 113 113 L -> V (IN REF. 4).
FT CONFLICT 297 297 MISSING (IN REF. 2 AND 3).
FT CONFLICT 357 357 S -> V (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 358 AA; 40681 MW; 6A99C521A5F6BB8 CRC64;
Query March 46.5%; Score 40; DB 1; Length 358;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 2 EWMNQHIDEERK 13
Db 72 EWMNVEHBEKQ 83

```

```

RESULT 15
CALX_SCHPO
ID CALX_SCHPO STANDARD; PRT; 560 AA.
AC P365B1;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Calnexin homolog precursor.
GN CAL1 OR CNX1 OR SPAC3C7.11C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95347333; PubMed=7621821;
RA Parlati F., Dignard D., Bergeron J.J.M., Thomas D.Y.;
RT "The calnexin homologue cnx1+ in Schizosaccharomyces pombe, is an
essential gene which can be complemented by its soluble ER domain.";
RL EMBO J. 14:3064-3072(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181485; PubMed=7876257;
RA Jamatpour M., Rokeach L.A.;
RT "The Schizosaccharomyces pombe homologue of the chaperone calnexin is
essential for viability.";
RL J. Biol. Chem. 270:4845-4853(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellgett T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Weller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Calcium-binding protein that interacts with newly
synthesized glycoproteins in the endoplasmic reticulum. It may act
in assisting protein assembly and/or in the retention within the
ER of unassembled protein subunits. It seems to play a major role
in the quality control apparatus of the ER by the retention of
incorrectly folded proteins.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
reticulum.
CC -!- SIMILARITY: Belongs to the calreticulin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL; M98799; AAA79757.1; -.  
 DR EMBL; U03389; AAA58631.1; -.  
 DR EMBL; Z93568; CAB16741.1; -.  
 DR PIR; S56142; S56142.  
 DR GeneDB; SPombe; SPAC3C7.11c; -.  
 DR InterPro; IPR000903; Calret\_calnex\_P.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR008985; Cona\_like\_lec\_g1.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR ProDom; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 2.  
 KM Calcium-binding; Endoplasmic reticulum; Transmembrane; Signal;  
 KM Repeat; Chaperone; Glycoprotein.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 560 CALNEXIN HOMOLOG.  
 FT DOMAIN 23 489 LUMENAL (POTENTIAL).  
 FT TRANSMEM 490 512 POTENTIAL.  
 FT DOMAIN 513 560 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 244 310 4 X APPROXIMATE REPEATS.  
 FT REPEAT 244 255 1-1.  
 FT REPEAT 261 272 1-2.  
 FT REPEAT 280 291 1-3.  
 FT REPEAT 299 310 1-4.  
 FT DOMAIN 314 371 4 X APPROXIMATE REPEATS.  
 FT REPEAT 314 324 2-1.  
 FT REPEAT 333 343 2-2.  
 FT REPEAT 347 357 2-3.  
 FT REPEAT 361 371 2-4.  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 560 AA; 63466 MW; E3B2087333E9123 CRC64;

Query Match 46.5%; Score 40; DB 1; Length 560;  
 Best Local Similarity 66.7%; Pred. No. 62;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 MHIDEKNE 15  
 ||:|||||:  
 Db 443 MHVDESNQ 451

Search completed: July 29, 2004, 11:35:22  
 Job time : 7.66667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:31:25 ; Search time 10.333 Seconds  
(Without alignments)  
139.633 Million cell updates/sec

Title: US-10-040-128-1

Perfect score: 86

Sequence: 1 DEIWNQMTIDEKNE 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	84.9	784	1	IJHUC5
2	50	58.1	77	2	A11306
3	49	57.0	77	2	A11678
4	47	54.7	712	2	T16338
5	45	52.3	391	2	C69673
6	45	52.3	2123	2	S55089
7	44	51.2	558	2	T29503
8	43	50.0	367	2	P97010
9	43	50.0	435	2	C64941
10	43	50.0	444	2	C89768
11	42	48.8	524	2	T20872
12	42	48.8	549	2	S49446
13	41	47.7	201	2	B97339
14	41	47.7	271	2	A81282
15	41	47.7	271	2	A81282
16	41	47.7	380	1	E69541
17	41	47.7	467	2	T10854
18	41	47.7	519	2	T33616
19	41	47.7	532	2	G90607
20	41	47.7	2241	2	T20971
21	41	47.7	2261	2	T20978
22	40	46.5	204	2	T05677
23	40	46.5	280	2	F81984
24	40	46.5	280	2	H81038
25	40	46.5	355	2	T06122
26	40	46.5	455	2	T40940
27	40	46.5	459	2	T29945
28	40	46.5	460	2	T29946
29	40	46.5	560	2	S56142

30	40	46.5	649	2	T01882	hypothetical prote
31	40	46.5	662	2	T18233	probable transcrip
32	40	46.5	665	2	T04290	hypothetical prote
33	40	46.5	844	2	T43112	hypothetical prote
34	40	46.5	866	2	T10587	serine/threonine-s
35	39.5	45.9	859	2	T01672	envelope polyprote
36	39.5	45.9	860	2	T38813	hypothetical prote
37	39.5	45.9	968	2	AD0514	probable ATP-depen
38	39	45.3	129	2	F72261	hypothetical prote
39	39	45.3	141	2	T51642	probable transcrip
40	39	45.3	146	1	HBOWP	hemoglobin beta ch
41	39	45.3	186	2	S35057	NADH2 dehydrogenas
42	39	45.3	231	2	T23136	hypothetical prote
43	39	45.3	239	2	G72455	probable glucose-1
44	39	45.3	270	2	F84255	hypothetical prote
45	39	45.3	318	2	T29785	hypothetical prote

ALIGNMENTS

RESULT 1

IJHUC5

cadherin 5 precursor - human

N/Alternate names: 7B4 antigen; cadherin, endothelial-specific; VE-cadherin

C/Species: Homo sapiens (man)

C/Date: 30-Jun-1993 #sequence revision 13-Sep-1996 #ext\_change 22-Jun-1999

C/Accession: S49893; S24305; A43418

R/Breviario, F.; Cavada, L.; Corda, M.; Martin-Padura, I.; Golay, J.; Introna, M.; Lamp

submitted to the EMBL Data Library, June 1994

A/Description: Molecular and functional properties of VE-cadherin (7B4/cadherin-5) a nov

A/Reference number: S49893

A/Accession: S49893

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-784 <BRE>

A/Cross-references: EMBL:X79981; NID:G599833; PIDN:CA456306.1; PID:G599834

R/Suzuki, S.; Sano, K.; Tanihara, H.

Cell Regul. 2, 261-270, 1991

A/Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous t

A/Reference number: S24305; MUID:91283540; PMID:2059658

A/Accession: S24305

A/Molecule type: mRNA

A/Residues: 5-516, 'I', 518-784 <SUZ>

A/Cross-references: EMBL:X59796; NID:G639976; PIDN:CA42468.1; PID:G29593

R/Lampugnani, M.G.; Resnati, M.; Raiteri, M.; Pigott, R.; Pissacane, A.; Hohen, G.; Ruco,

J. Cell Biol. 118, 1511-1522, 1992

A/Title: A novel endothelial-specific membrane protein is a marker of cell-cell contacts

A/Reference number: A43418; MUID:92394977; PMID:1522121

A/Accession: A43418

A/Molecule type: protein

A/Residues: 48-60, 'X', 62, 'X', 64, 108-116, 'X', 118-123, 237-238, 'X', 240, 'X', 242-252, 'X', 254-5

A/Experimental source: cultured endothelial cells

A/Note: sequence extracted from NCBI Backbone (NCBI:P:113040, NCBI:P:113045, NCBI:P:113047,

C/Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to

C/Genetics:

A/Gene: GDB:CDH5

A/Cross-references: GDB:134230; OMIM:601120

A/Map position: 16q22.1-16q22.1

C/Superfamily: cadherin; cadherin repeat homology

C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane prot

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-47/Domain: propeptide #status predicted <PRO>

F:48-784/Product: cadherin 5 #status predicted <MAT>

F:48-593/Domain: extracellular #status predicted <EXT>

F:50-151/Domain: cadherin repeat homology <CR1>

F:154-258/Domain: cadherin repeat homology <CR2>

F:261-372/Domain: cadherin repeat homology <CR3>

F:375-479/Domain: cadherin repeat homology <CR4>

F:481-587/Domain: cadherin repeat homology <CR5>

F:594-620/Domain: transmembrane #status predicted <TM>

F:621-784/Domain: intracellular #status predicted <INT>

F:736-753/Region: serine-rich



probable acetyl-CoA carboxylase (EC 6.4.1.2) HPA1 - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: protein YM8261.01c; protein YM825.08c; protein YMR207c  
 C:Species: *Saccharomyces cerevisiae*  
 C>Date: 08-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 11-Jan-2002  
 C/Accession: S55089; S41802; S59447  
 R:Dedman, K.; Brown, D.; Bowman, S.  
 submitted to the EMBL Data Library, June 1995  
 A:Reference number: S55089  
 A:Accession: S55089  
 A:Molecule type: DNA  
 A:Residues: 1-833 <DED>  
 A:Cross-references: EMBL:Z49809; MIPS:YMR207c; NID:g954459; PIDN:CAA89922.1; PID:g132600  
 A:Experimental source: strain AB972  
 A>Note: the published sequence extends beyond the amino end  
 R:Kearney, S.E.  
 submitted to the EMBL Data Library, April 1993  
 A:Description: Identification of an *Saccharomyces cerevisiae* gene closely related to FAS  
 A:Reference number: S41802  
 A:Accession: S41802  
 A:Molecule type: DNA  
 A:Residues: 1-510, 'U', 512-799 <KEA>  
 A:Cross-references: EMBL:Z22558; NID:g296212; PIDN:CAA80280.1; PID:g388250  
 A>Note: the published sequence extends beyond the amino end  
 R:Odell, C.; Bowman, S.  
 submitted to the EMBL Data Library, March 1995  
 A:Reference number: S59441  
 A:Accession: S59447  
 A:Molecule type: DNA  
 A:Residues: 812-2123 <ODE>  
 A:Cross-references: EMBL:Z48755; MIPS:YMR207c; NID:g736296; PIDN:CAA86647.1; PID:g763183  
 A:Experimental source: strain AB972  
 C:Genetics:  
 A:Gene: SGD:HPA1  
 A:Cross-references: MIPS:YMR207c; SGD:S0004820  
 A:Map position: 13R  
 C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-  
 C:Keywords: biotin binding; ligase  
 F:1-487/Domain: biotin carboxylase homology <ECH>  
 F:615-687/Domain: lipoyl/biotin-binding homology <LPB>  
 F:654/Binding site: biotin (lys) (covalent) #status predicted

Query Match 52.3%; Score 45; DB 2; Length 2123;  
 Best Local Similarity 53.8%; Pred. No. 96;  
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 DEINQMHIDEK 13  
 Db 1098 DEVEQIHPEER 1110

RESULT 7  
 T29503  
 serie C-palmitoyltransferase (EC 2.3.1.50) F43H9.2 [similarity] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
 C/Accession: T29503  
 R:Mu, X.; Le, T.  
 submitted to the EMBL Data Library, February 1996  
 A:Description: The sequence of *C. elegans* cosmid F43H9.  
 A:Reference number: Z20629  
 A:Accession: T29503  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-558 <WUX>  
 A:Cross-references: EMBL:U50307; PIDN:AAA92303.1; CESP:F43H9.2  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:F43H9.2  
 A:Insertions: 16/2; 71/3; 115/3; 289/1; 528/3  
 C:Superfamily: serie C-palmitoyltransferase chain LC82; glycine C-acetyltransferase hom  
 C:Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate  
 F:18-531/Domain: glycine C-acetyltransferase homology <GGA>  
 F:384/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 51.2%; Score 44; DB 2; Length 558;  
 Best Local Similarity 58.3%; Pred. No. 31;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 WNOQHIDEKNE 15  
 Db 29 WNOEHDDDEEE 40

RESULT 8  
 F97010  
 fusion, chorismate mutase and shikimate 5-dehydrogenase [imported] - *Clostridium acetobu*  
 C:Species: *Clostridium acetobutylicum*  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C/Accession: F97010  
 R:Mollig, U.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: F97010  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-367 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK78873.1; PID:g15023795; GSPDB:GN00168  
 A:Experimental source: *Clostridium acetobutylicum* ATCC824  
 C:Genetics:  
 A:Gene: CAC0897

Query Match 50.0%; Score 43; DB 2; Length 367;  
 Best Local Similarity 63.6%; Pred. No. 28;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DEINQMHIDE 11  
 Db 341 EIMWIDISIDE 351

RESULT 9  
 C84941  
 histidinol dehydrogenase (EC 1.1.1.23) [imported] - *Buchnera* sp. (strain APS)  
 C:Species: *Buchnera* sp.  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C/Accession: C84941  
 R:Shigenobu, S.; Matanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
 Nature 407, 81-86, 2000  
 A>Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. A  
 A:Reference number: AB4930; MUID:20445173; PMID:10993077  
 A:Accession: C84941  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-435 <STO>  
 A:Cross-references: GB:AP000398; GSPDB:GN00144  
 A:Experimental source: strain APS  
 C:Genetics:  
 A:Gene: hsdJ; BU100  
 C:Superfamily: histidinol dehydrogenase; histidinol dehydrogenase homology  
 C:Keywords: oxidoreductase

Query Match 50.0%; Score 43; DB 2; Length 435;  
 Best Local Similarity 60.0%; Pred. No. 34;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WNOQHIDEK 13  
 Db 10 WNKJHPDEOK 19

RESULT 10  
 C89768  
 conserved hypothetical protein SA0083 [imported] - *Staphylococcus aureus* (strain N315)  
 C:Species: *Staphylococcus aureus*

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C/Accession: C89768  
R/Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C/Species: *Shibui, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.*  
Lancet 357, 1225-1240, 2001  
A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A/Reference number: A89758; MUID:21311952; PMID:11418146  
A/Accession: C89768  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-444 <KIR>  
A/Cross-references: GB:BA000018; PID:G13700003; PIDN:BA841302.1; GSPDB:GN00149  
A/Experimental source: strain N315  
C/Genetics:  
A/Gene: SA0083

Query Match 50.0%; Score 43; DB 2; Length 444;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DEINQNHIDEKN 14  
|||:|:|:|:  
Db 372 DEWMNGHLDQAVN 385

RESULT 11  
T20872  
hypothetical protein F26H9.1 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T20872; T21443  
R/Sims, M.  
submitted to the EMBL Data Library, June 1996  
A/Reference number: Z19337  
A/Accession: T20872  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-524 <WII>  
A/Cross-references: EMBL:Z75535; PIDN:CAA99829.1; GSPDB:GN00019; CESP:F26H9.1  
A/Experimental source: clone F14B4  
R/Baynes, C.  
submitted to the EMBL Data Library, November 1996  
A/Reference number: Z19422  
A/Accession: T21443  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-524 <W12>  
A/Cross-references: EMBL:Z81516; PIDN:CAB04206.1; GSPDB:GN00019; CESP:F26H9.1  
A/Experimental source: clone F26H9  
C/Genetics:  
A/Gene: CESP:F26H9.1  
A/Map position: 1  
A/Introns: 49/2; 99/1; 136/1; 186/1; 228/3; 287/3; 346/1; 403/1; 449/3

Query Match 48.8%; Score 42; DB 2; Length 524;  
Best Local Similarity 50.0%; Pred. No. 60;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 WNMNHIDEKN 15  
||:|:|:|:  
Db 488 WHQRNIDEHRNQ 499

RESULT 12  
S49446  
RING-finger protein - *Lotus japonicus*  
C/Species: *Lotus japonicus*  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 23-Mar-2001  
C/Accession: S49446  
R/Schauser, U.; Borg, S.; Poulsen, C.  
submitted to the EMBL Data Library, August 1994  
A/Description: Pzf, a gene isolated from a *Lotus japonicus* nodule cDNA library, encodes

A/Reference number: S49445  
A/Accession: S49446  
A/Molecule type: mRNA  
A/Residues: 1-549 <SCH>  
A/Cross-references: EMBL:Z6750; NID:G558544; PIDN:CAA85321.1; PID:G558545  
A/Experimental source: strain Gifu B-129  
C/Superfamily: RING finger homology  
C/Keywords: DNA binding; zinc  
F/492-543/Domain: RING finger homology <RNG>

Query Match 48.8%; Score 42; DB 2; Length 549;  
Best Local Similarity 54.5%; Pred. No. 64;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEINQNHIDE 11  
||:|:|:|:  
Db 18 DQNMNMHTDQ 28

RESULT 13  
E97339  
HAD superfamily hydrolase [imported] - *Clostridium acetobutylicum*  
C/Species: *Clostridium acetobutylicum*  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C/Accession: E97339  
R/Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*  
A/Reference number: A96900; MUID:2135325; PMID:21359325  
A/Accession: E97339  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-201 <KUR>  
A/Cross-references: GB:AE001437; PIDN:AAK81504.1; PID:G15026677; GSPDB:GN00168  
A/Experimental source: *Clostridium acetobutylicum* ATCC824  
C/Genetics:  
A/Gene: CAC3581

Query Match 47.7%; Score 41; DB 2; Length 201;  
Best Local Similarity 54.5%; Pred. No. 30;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EWMNMHVPE 88  
|||:|:|:  
Db 78 EWMNMHVPE 88

RESULT 14  
AB1292  
amino acid ABC transporter (binding protein) homolog lmo1738 [imported] - *Listeria monoc.*  
C/Species: *Listeria monocytogenes*  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C/Accession: AB1292  
R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, J.; Dominguez-Bernal, G.; Duchaud, B.; Durand, L.; Dussurget, O.; Entian, K.D.; Fstli, H. D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitouram, A.; Maok, C.; Schuener, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, A./Title: Comparative genomics of *Listeria* species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AB1292  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-271 <GLA>  
A/Cross-references: GB:NC\_003210; PIDN:CAC99816.1; PID:G16411192; GSPDB:GN00177  
A/Experimental source: strain EGD-e  
C/Genetics:  
A/Gene: lmo1738  
C/Superfamily: Lysine-arginine-ornithine-binding protein

Query Match 47.7%; Score 41; DB 2; Length 271;

Best Local Similarity 40.0%; Pred. No. 42;  
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DEINQMHIIDEKNE 15  
|:::|:::|:::  
Db 198 DEVYARYYIDKQKNK 212

RESULT 15

AH1663  
amino acid ABC transporter (binding protein) homolog lin1849 [imported] - *Listeria innocua*  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AH1663  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tixeront, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MIMD:21537279; PMID:11679669  
A:Accession: AH1663  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-271 <GIA>  
A:Cross-references: GB:AL592022; PIDN:CA697080.1; PID:gl6414351; GSPDB:GN00178  
A:Experimental source: strain C10P11262  
C:Genetics:  
A:Gene: lin1849  
C:Superfamily: Lysine-arginine-ornithine-binding protein

Query Match 47.7%; Score 41; DB 2; Length 271;  
Best Local Similarity 40.0%; Pred. No. 42;  
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
QY 1 DEINQMHIIDEKNE 15  
|:::|:::|:::  
Db 198 DEVYARYYIDKQKNK 212

Search completed: July 29, 2004, 11:37:49  
Job time : 11.3333 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:29:40 ; Search time 46.667 Seconds

(without alignments)  
90.819 Million cell updates/sec

Title: US-10-040-128-1

Sequence: 1 DEIMQMHIKXNE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	15	AAU08777	AAU08777 VE-cadher
2	78	90.7	15	AAU08778	AAU08778 VE-cadher
3	78	90.7	103	AAU08786	AAU08786 Murine VE
4	73	84.9	15	AAU08779	AAU08779 VE-cadher
5	73	84.9	104	AAU08785	AAU08785 Human VE-
6	73	84.9	108	AAV64578	AAV64578 Nonclasti
7	73	84.9	566	AAU11276	AAU11276 Human rec
8	73	84.9	594	AAU11279	AAU11279 Human rec
9	73	84.9	780	AAW25634	AAW25634 Human cad
10	73	84.9	780	AAW13130	AAW13130 Full leng
11	73	84.9	784	ABB81477	ABB81477 Human VE-
12	73	84.9	784	ABU03507	ABU03507 Angiogene
13	73	84.9	784	ABU56570	ABU56570 Lung can
14	73	84.9	784	ADA03363	ADA03363 Human vas
15	50	58.1	77	ABB49545	ABB49545 Listeria
16	46	53.5	183	ABB58822	ABB58822 Breast an
17	46	53.5	192	ADC26203	ADC26203 Human NOV
18	46	53.5	192	ADC10034	ADC10034 Human NOV
19	46	53.5	192	ADC10028	ADC10028 Human NOV
20	46	53.5	193	AAU23017	AAU23017 Novel hum
21	46	53.5	193	ABH10365	ABH10365 Human CDN
22	46	53.5	193	AAU18501	AAU18501 Human end
23	46	53.5	193	ABP66952	ABP66952 Human pol
24	46	53.5	1032	ABR39808	ABR39808 Human SCA
25	45	52.3	418	ABU46642	ABU46642 Protein e

26	45	52.3	420	5	ABP28693	ABP28693 Streptoco
27	43	50.0	105	5	ABP34429	ABP34429 Human red
28	43	50.0	132	4	AAO05753	AAO05753 Human pol
29	43	50.0	367	6	ABU23606	ABU23606 Protein e
30	43	50.0	444	6	ABM73085	ABM73085 Staphyloc
31	42	48.8	75	2	AAI11833	AAI11833 Human 5'
32	42	48.8	145	4	ABG16313	ABG16313 Novel hum
33	42	48.8	254	4	AAI14054	AAI14054 Novel hum
34	42	48.8	254	4	ABH32999	ABH32999 Peptide #
35	42	48.8	254	4	AAW26460	AAW26460 Peptide #
36	42	48.8	254	4	ABE27828	ABE27828 Human pep
37	42	48.8	254	4	ABH18472	ABH18472 Protein #
38	42	48.8	254	4	AAW66183	AAW66183 Human bon
39	42	48.8	254	4	AAW53797	AAW53797 Human btx
40	42	48.8	254	4	ABG47851	ABG47851 Human liv
41	42	48.8	254	4	AAW01794	AAW01794 Peptide #
42	42	48.8	254	5	ABG35833	ABG35833 Human pep
43	42	48.8	679	4	ABG10004	ABG10004 Novel hum
44	42	48.8	1039	3	AAV58837	AAV58837 Soybean c
45	42	48.8	1113	4	ABG10005	ABG10005 Novel hum

#### ALIGNMENTS

RESULT 1	AAU08777	standard, peptide; 15 AA.
ID	AAU08777	standard, peptide; 15 AA.
XX	AAU08777;	
DT	27-FEB-2002	(first entry)
XX	VE-cadherin N-terminal domain 1 antibody binding peptide #1.	
DE	VE-cadherin, antibody binding peptide; angiogenesis; neoplastic disease;	
XX	VE-cadherin mediated adherens junction formation; autoimmune disease;	
KW	paracellular permeability; solid tumour; collagenous vascular disease;	
KW	rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;	
KW	retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;	
KW	carcinoma; sarcoma; adenocarcinoma; adenosarcoma; adenoma; gene therapy;	
KW	leukaemic tumour; lymphoid tumours; cell proliferation disorder;	
KW	vacuolarisation; blood vessel proliferation disorder; fibrotic disorder;	
KW	age-related muscular degeneration; tumour neovascularisation;	
KW	VE-cadherin N-terminal domain 1.	
OS	Mammalia.	
XX	WO200175109-A2.	
XX	11-OCT-2001.	
PD	30-MAR-2001; 2001WO-US010505.	
XX	31-MAR-2000; 2000US-00540967.	
PR	(IMCL-) IMCLONE SYSTEMS INC.	
XX	Liiao F, Hicklin DJ, Bohlen P;	
PI	WPI; 2001-655988/75.	
XX	New antibody antagonists of VE-cadherin, which does not adversely affect	
PT	vascular permeability, useful for inhibiting angiogenesis or tumor	
PT	metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.	
PS	Claim 1; Page 32; 44pp; English.	
XX	The method relates to an antibody or an antibody fragment, capable of	
CC	specifically binding to a VE-cadherin, and of inhibiting VE-cadherin	
CC	mediated adherens junction formation in vitro, but does not exert any	
CC	significant or substantial effect on paracellular permeability in vitro.	
CC	The antibody or antibody fragment is capable of specifically binding to a	

CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 CC adenosarcomas, adenomas, leukaemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vasculature in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
 CC domain 1 antibody binding peptide

XX Sequence 15 AA;

Query Match 100.0%; Score 86; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.9e-07; Mismatches 0; Gaps 0;

Matches 15; Conservative 0; Indels 0; Gaps 0;

OY 1 DE1WNQNHIDEKNE 15  
 Db 1 DE1WNQNHIDEKNE 15

RESULT 2

AAU08778 standard; peptide; 15 AA.

XX AAU08778;

XX 27-FEB-2002 (first entry)

XX VE-cadherin N-terminal domain 1 antibody binding peptide #2.

XX VE-cadherin; antibody binding peptide; angiogenesis; neoplastic disease;  
 XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
 XX paracellular permeability; solid tumour; collagenous vascular disease;  
 XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
 XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
 XX carcinoma; sarcoma; adenocarcinoma; adenoma; gene therapy;  
 XX leukaemic tumour; lymphoid tumours; cell proliferation disorder;  
 XX vascularisation; blood vessel proliferation disorder; fibrotic disorder;  
 XX age-related muscular degeneration; tumour neovascularisation;  
 XX VE-cadherin N-terminal domain 1.

XX Mammalia.

XX WO200175109-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010505.

XX 31-MAR-2000; 2000US-00540967.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX Liao F, Hicklin DJ, Bohlen P;

XX WPI; 2001-656988/75.

XX New antibody antagonists of VE-cadherin, which does not adversely affect  
 XX PT vascular permeability, useful for inhibiting angiogenesis or tumour  
 XX PT metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.

XX Claim 1; Page 32; 44pp; English.

XX The method relates to an antibody or an antibody fragment, capable of

CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 CC mediated adherens junction formation in vitro, but does not exert any  
 CC significant or substantial effect on paracellular permeability in vitro.  
 CC The antibody or antibody fragment is capable of specifically binding to a  
 CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 CC adenosarcomas, adenomas, leukaemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vasculature in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
 CC domain 1 antibody binding peptide

XX Sequence 15 AA;

Query Match 90.7%; Score 78; DB 4; Length 15;

Best Local Similarity 93.3%; Pred. No. 1.3e-05; Mismatches 1; Gaps 0;

Matches 14; Conservative 0; Indels 1; Gaps 0;

OY 1 DE1WNQNHIDEKNE 15  
 Db 1 DE1WNQNHIDEKNE 15

RESULT 3

AAU08786 standard; protein; 103 AA.

XX AAU08786;

XX 27-FEB-2002 (first entry)

XX Murine VE-cadherin N-terminal domain 1 polypeptide.

XX VE-cadherin; angiogenesis; neoplastic disease;  
 XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
 XX paracellular permeability; solid tumour; collagenous vascular disease;  
 XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
 XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
 XX carcinoma; sarcoma; adenocarcinoma; adenoma; gene therapy;  
 XX leukaemic tumour; lymphoid tumours; cell proliferation disorder;  
 XX vascularisation; blood vessel proliferation disorder; fibrotic disorder;  
 XX age-related muscular degeneration; tumour neovascularisation; mouse;  
 XX VE-cadherin N-terminal domain 1.

XX Mus sp.

XX WO200175109-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010505.

XX 31-MAR-2000; 2000US-00540967.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX Liao F, Hicklin DJ, Bohlen P;

XX WPI; 2001-656988/75.

XX New antibody antagonists of VE-cadherin, which does not adversely affect  
 XX PT vascular permeability, useful for inhibiting angiogenesis or tumour  
 XX PT metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.





XX	Liao F, Hicklin DJ, Bohlen P;
PI	WPI; 2001-656988/75.
XX	
PT	New antibody antagonists of VE-cadherin, which does not adversely affect
PT	vascular permeability, useful for inhibiting angiogenesis or tumor
PT	metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.
XX	
PS	Example 2; Fig 2; 44pp; English.
XX	
CC	The method relates to an antibody or an antibody fragment, capable of
CC	specifically binding to a VE-cadherin, and of inhibiting VE-cadherin
CC	mediated adherens junction formation in vitro, but does not exert any
CC	significant or substantial effect on paracellular permeability in vitro.
CC	The antibody or antibody fragment is capable of specifically binding to a
CC	site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of
CC	a VE-cadherin. The antibody or antibody fragment is useful for inhibiting
CC	angiogenesis (e.g. angiogenesis that is associated with a neoplastic
CC	disease), a solid tumour, an autoimmune disease, collagenous vascular
CC	disease, rheumatoid arthritis, an ophthalmological condition, diabetic
CC	retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour
CC	metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,
CC	adenosarcomas, adenomas, leukaemic tumours or lymphoid tumours). The
CC	composition or antibody is also useful for treating a cell proliferation
CC	disorder associated with vascularisation (e.g. blood vessel proliferation
CC	disorders, fibrotic disorders, angiogenesis, tumour growth, tumour
CC	metastasis, rheumatoid arthritis or age-related muscular degeneration).
CC	These may also be used for reducing or inhibiting tumour vasculature in a
CC	mammal. The nucleic acid that encodes the antibody or is useful in gene
CC	therapy, particularly for inhibiting angiogenesis or tumour
CC	neovascularisation. This sequence represents a human VE-cadherin N-
CC	terminal domain 1 polypeptide
SQ	Sequence 104 AA;
OY	Query Match                      84.9%; Score 73; DB 4; Length 104;
Dn	Best Local Similarity    92.9%; Pred. No. 0.00068;
	Matches    13; Conservative    0; Mismatches    1; Indels    0; Gaps    0
1	DEINQMHIIDEKN 14
1	DWINQMHIIDEKN 14
RESULT 6	
AA64578	
ID	AA64578 standard; peptide; 108 AA.
AC	
XX	AA64578;
DT	
XX	02-MAR-2000 (first entry)
DE	
XX	Nonclassical cadherin extracellular domain SEQ ID NO:6.
XX	
KM	Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KM	inhibition; cadherin extracellular domain; cell adhesion recognition;
KM	OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KM	cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KM	cadherin related neuronal receptor; LI-cadherin; protocadherin;
KM	desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KM	rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KM	neurological disease.
OS	Mammaria.
PN	WO957149-A2.
XX	
PD	11-NOV-1999.
XX	
PF	05-MAY-1999; 99MO-CAN00363.
XX	
PR	05-MAY-1998; 98US-00073040.

[illegible]

```

XX 04-MAY-2000; 2000GB-00010630.
XX (UYMA-) UNIV WALES COLLEGE OF MEDICINE.
XX Jiang WG;
XX WPI; 2002-066521/09.
XX N-PSDB; AAS16965.
XX Novel nucleic acid for inhibition of angiogenesis and prophylaxis or
XX treatment of cancer, comprises both an angiogenic factor antagonist and a
XX vascular endothelial structure regulator.
XX Disclosure; Fig 2; 66pp; English.
XX The invention relates to recombinant nucleic acid sequences that encode
XX both an angiogenic factor antagonist and a vascular endothelial structure
XX regulator, collectively referred to as a KV protein. KV DNA and proteins,
XX and also cells and compositions containing the sequences, are useful in
XX treating or preventing cancer or angiogenesis in mammals, in particular
XX humans. KV proteins are also useful for regulation of the development of
XX blood vessels and their formation, in the vascular endothelium and/or
XX tumour. The KVE702 gene and its fragments are useful in transfection of
XX human epithelial cells and to generate products suitable for angiogenesis
XX intervention. This sequence represents the human KVE702 protein
XX
SQ Sequence 566 AA;
XX
Query Match      84.9%; Score 73; DB 5; Length 566;
Best Local Similarity 92.9%; Pred. No. 0.0044;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DEINQMHIIDEKN 14
   | |||||
Db 474 DWIMQMHIIDEKN 487

RESULT 8
AAU11279
ID AAU11279 standard; protein; 594 AA.
XX
AC AAU11279;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human recombinant U37 polypeptide.
XX
KW Human; KV protein; KVE702; angiogenic factor antagonist; angiogenesis;
XX vascular endothelial structure regulator; cancer; vascular endothelium;
XX tumour; blood vessel; epithelial cell; cytostatic; gene therapy; U37.
XX
OS Homo sapiens.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 518 /note= "Encoded by ATC"
XX
PN WO200183562-A2.
XX
PD 08-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-GB001956.
XX
PR 04-MAY-2000; 2000GB-00010630.
XX
XX (UYMA-) UNIV WALES COLLEGE OF MEDICINE.
XX Jiang WG;
XX WPI; 2002-066521/09.
XX N-PSDB; AAS16975.

```

```

XX Novel nucleic acid for inhibition of angiogenesis and prophylaxis or
XX treatment of cancer, comprises both an angiogenic factor antagonist and a
XX vascular endothelial structure regulator.
XX Disclosure; Fig 15; 66pp; English.
XX The invention relates to recombinant nucleic acid sequences that encode
XX both an angiogenic factor antagonist and a vascular endothelial structure
XX regulator, collectively referred to as a KV protein. KV DNA and proteins,
XX and also cells and compositions containing the sequences, are useful in
XX treating or preventing cancer or angiogenesis in mammals, in particular
XX humans. KV proteins are also useful for regulation of the development of
XX blood vessels and their formation, in the vascular endothelium and/or
XX tumour. The KVE702 gene and its fragments are useful in transfection of
XX human epithelial cells and to generate products suitable for angiogenesis
XX intervention. This sequence represents the human U37 protein used in
XX methods of the invention
XX
SQ Sequence 594 AA;
XX
Query Match      84.9%; Score 73; DB 5; Length 594;
Best Local Similarity 92.9%; Pred. No. 0.0047;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DEINQMHIIDEKN 14
   | |||||
Db 49 DWIMQMHIIDEKN 62

RESULT 9
AAM25634
ID AAM25634 standard; protein; 780 AA.
XX
AC AAM25634;
XX
DT 25-MAR-2003 (revised)
DT 03-NOV-1997 (first entry)
XX
DE Human cadherin-5.
XX
KW Human; cadherin; rat; calcium-dependent cell adhesion protein;
XX superfamily; cytoskeleton; eatenin; cancer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 622 /note= "encoded by AAG"
XX
PN US5646250-A.
XX
PD 08-JUL-1997.
XX
PF 01-NOV-1994; 94US-00332638.
XX
PR 17-APR-1992; 92US-00872643.
XX
PR 19-APR-1993; 93US-00049460.
XX
XX (DOHR-) DOHERY EYE INST.
XX
XX Suzuki S;
XX WPI; 1997-362997/33.
XX N-PSDB; AAT85401.
XX
XX Human and rat cadherin polypeptide(s) - mediate cell-cell adhesion.
XX Claim 1; Col 69-72; 56pp; English.
XX
XX This sequence represents human cadherin-5. The invention specifically
XX provides details of human cadherin-5, -8, -11, -12 and -13, and rat
XX cadherin-8, -5, -11 and -13. Cadherins are calcium-dependent cell

```

CC adhesion proteins. They are glycosylated integral membrane proteins that  
 CC have an N-terminal extracellular domain that determines binding  
 CC specificity, a hydrophobic membrane spanning region and a C-terminal  
 CC cytoplasmic domain, which is highly conserved among members of the  
 CC superfamily. The C-terminal domain interacts with the cytoskeleton  
 CC through catenins and other cytoskeleton-associated proteins. The novel  
 CC cadherin proteins may be used in the analysis of the role of cadherins in  
 CC various cancers. Sequence analysis of the cadherin proteins also allows  
 CC investigation of the structure and function of cadherin. The cadherin  
 CC proteins may be isolated by using anti-cadherin antibodies. These  
 CC antibodies may also be used to modulate the activity of cadherin and to  
 CC determine the tissue specific distribution of cadherin proteins. Each  
 CC subclass of cadherins has a unique tissue distribution pattern. (Updated  
 CC on 25-MAR-2003 to correct PF field.)  
 CC  
 CC Sequence 780 AA;  
 CC  
 CC Query Match 84.9%; Score 73; DB 2; Length 780;  
 CC Best Local Similarity 92.9%; Pred. No. 0.0063;  
 CC Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 DEINQMHIIDEKN 14  
 CC | | | | | | | | | | | | | | | | | |  
 CC 44 DWIMQMHIIDEKN 57  
 CC  
 CC Db  
 CC  
 CC RESULT 10  
 CC AAM13130  
 CC ID AAM13130 standard; protein; 780 AA.  
 CC XX  
 CC AC AAM13130;  
 CC DT 25-MAR-2003 (revised)  
 CC DT 13-MAY-1997 (first entry)  
 CC XX  
 CC DE Full length human cadherin-5.  
 CC XX  
 CC XX Ca2+ dependent; cell adhesion protein; placental; cadherin; rat; brain;  
 CC KM human; antibody; purification; determination; tissue expression;  
 CC KW binding antagonist; calcium ion.  
 CC XX  
 CC OS Homo sapiens.  
 CC XX  
 CC XX US5597725-A.  
 CC PN 28-JAN-1997.  
 CC XX  
 CC PD 26-JAN-1994; 94US-00188228.  
 CC PF 17-APR-1992; 92US-00872643.  
 CC PR 19-APR-1993; 93US-00049460.  
 CC XX  
 CC PA (DOHE-) DOHENY EYE INST.  
 CC XX  
 CC PI Suzuki S;  
 CC XX  
 CC DR WPI; 1997-108328/10.  
 CC DR N-PSDB; AAT61921.  
 CC XX  
 CC PT Antibodies to cadherin proteins - useful as cadherin antagonists, etc.  
 CC XX  
 CC PS Claim 5; Col 75-78; 59pp; English.  
 CC XX  
 CC CC The present sequence is full length human cadherin-5, which is a Ca2+  
 CC dependent cell adhesion protein. The human cadherin cDNA was isolated  
 CC from a placental cDNA library, using probes based on homologous rat  
 CC cadherin cDNA. Antibodies or fragments that specifically bind the human  
 CC cadherin can be used to purify the cadherin, determine its tissue  
 CC expression and antagonise its ligand/antiligand binding activities.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 CC Sequence 780 AA;  
 CC  
 CC SQ

Query Match 84.9%; Score 73; DB 2; Length 780;  
 Best Local Similarity 92.9%; Pred. No. 0.0063;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DEINQMHIIDEKN 14  
 | | | | | | | | | | | | | | | | | |  
 44 DWIMQMHIIDEKN 57  
 Db  
 RESULT 11  
 ABB81477  
 ID ABB81477 standard; protein; 784 AA.  
 XX  
 AC ABB81477;  
 XX  
 DT 30-AUG-2002 (first entry)  
 XX  
 DE Human VE-cadherin protein SEQ ID NO:52.  
 XX  
 DE Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;  
 XX o-catenin; colon cancer.  
 KW  
 XX  
 OS Homo sapiens.  
 OS  
 PN US2002045591-A1.  
 PN  
 PD 18-APR-2002.  
 PD  
 PF 17-JUL-2001; 2001US-00905983.  
 PF  
 XX 26-MAY-1998; 98IL-00124650.  
 PR 26-MAY-1999; 99US-00318633.  
 PR  
 XX (GEIG/) GEIGER B.  
 PA (BENZ/) BEN-ZE'EV A.  
 PA (SADO/) SADO E.  
 XX  
 PI Geiger B, Ben-Ze'ev A, Sadot E;  
 XX  
 DR WPI; 2002-499105/53.  
 DR  
 DR N-PSDB; ABN89394.  
 DR  
 PT New construct encoding soluble cytoplasmic portion of cadherin including  
 PT beta catenin binding domain useful in treating cancer associated with  
 PT high beta-catenin activity e.g. colon cancer and melanoma.  
 PT  
 XX  
 XX Example 3; Page 56-58; 102pp; English.  
 PS  
 CC The present invention describes a pharmaceutical composition for treating  
 CC cancer associated with abnormally high beta-catenin activity. The  
 CC pharmaceutical composition comprises a gene therapy vehicle harbouring a  
 CC polynucleotide that contains: (i) a nucleotide sequence encoding a  
 CC soluble cytoplasmic portion of a cadherin which lacks a transmembrane  
 CC portion and an extracellular portion of the cadherin, and includes a beta  
 CC -catenin binding domain; and (b) an upstream promoter for directing  
 CC expression of the soluble cytoplasmic portion of the cadherin in a  
 CC mammalian cell. Also described is a pharmaceutical composition for  
 CC treating cancer associated with abnormally high activity levels of beta-  
 CC catenin comprising a gene therapy vehicle harbouring a polynucleotide  
 CC that contains: (a) a nucleotide sequence encoding an o-catenin; and (b)  
 CC an upstream promoter for directing expression of the o-catenin in a  
 CC mammalian cell. The pharmaceutical compositions have cytostatic activity  
 CC and can be used in the suppression of beta-catenin-mediated  
 CC transactivation. They can be used for treating cancers associated with  
 CC abnormally high activity levels of beta-catenin such as colon cancers and  
 CC melanomas, by reducing these high activity levels of beta-catenin in  
 CC mammalian cells. The present sequence represents human VE-cadherin which  
 CC is used in the exemplification of the present invention  
 CC  
 CC Sequence 784 AA;  
 CC  
 CC SQ  
 CC Query Match 84.9%; Score 73; DB 5; Length 784;  
 CC Best Local Similarity 92.9%; Pred. No. 0.0064;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 DEIMNQMHIDEKN 14  
| | | | | | | | | | | | | | | | | |  
Db 48 DWIMNQMHIDEKN 61

RESULT 12  
ABU03507  
ID ABU03507 standard; protein; 784 AA.  
XX  
AC ABU03507;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Angiogenesis-associated human protein sequence #52.  
XX  
KW Human; angiogenesis-associated transcript; angiogenesis;  
KW angiogenesis-associated disease; cancer; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200279492-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 14-FEB-2002; 2002WO-US004915.  
XX  
PR 14-FEB-2001; 2001US-00794356.  
PR 22-FEB-2001; 2001US-00791390.  
PR 19-APR-2001; 2001US-0285475P.  
PR 03-AUG-2001; 2001US-0310025P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334244P.  
XX  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
PI Murray R, Glyme R, Watson SR, Aziz N;  
DR WPI; 2003-040681/03.  
DR N-PSDB; ABX08791.  
XX  
XX  
PT Detecting angiogenesis-associated transcript in a cell for diagnosing and  
PT treating cancer by contacting a sample with a polynucleotide that  
PT exhibits changes in expression level as a function of time in tissue  
XX undergoing angiogenesis.  
XX  
PS Example 2; Page 230; 291pp; English.  
XX  
CC The present invention relates to methods and compositions for detecting  
CC an angiogenesis-associated transcript in a cell in a patient. The method  
CC involves contacting a biological sample from the patient with a  
CC polynucleotide that selectively hybridizes to a sequence at least 80%  
CC identical to any of the angiogenesis-associated human polynucleotide  
CC sequences given in the specification. These angiogenesis-associated  
CC polynucleotide sequences comprise genes that exhibit changes in  
CC expression levels as a function of time in tissue undergoing  
CC angiogenesis. The method and the polynucleotide sequences of the  
CC invention are useful for diagnosing and treating angiogenesis and  
CC angiogenesis-associated diseases e.g. cancer. The polynucleotide  
CC sequences are also useful in the gene therapy of such disorders. The  
CC angiogenesis-associated proteins encoded by the polynucleotide sequences  
CC are useful as a vaccine for therapeutic and prophylactic immunisation.  
CC ABU03456-ABU03569 represent angiogenesis-associated protein sequences  
XX  
SQ Sequence 784 AA;

Query Match 84.9%; Score 73; DB 6; Length 784;  
Best Local Similarity 92.9%; Pred. No. 0.0064;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEIMNQMHIDEKN 14  
| | | | | | | | | | | | | | | | | |

Db 48 DWIMNQMHIDEKN 61

RESULT 13  
ABU56570  
ID ABU56570 standard; protein; 784 AA.  
XX  
AC ABU56570;  
XX  
DT 02-APR-2003 (first entry)  
XX  
DE Lung cancer-associated polypeptide #163.  
XX  
KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
XX  
OS Unidentified.  
XX  
PN WO200286443-A2.  
XX  
PD 31-OCT-2002.  
XX  
PF 18-APR-2002; 2002WO-US012476.  
XX  
PR 18-APR-2001; 2001US-0284770P.  
PR 10-MAY-2001; 2001US-0290492P.  
PR 09-NOV-2001; 2001US-0339245P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334370P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
PI Aziz N, Murray R;  
DR WPI; 2003-093161/08.  
DR N-PSDB; ABX76299.  
XX  
XX  
PT Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the  
PT patient with a polynucleotide that exhibits increased or decreased  
XX expression in lung cancer.  
XX  
PS Claim 27; Page 314; 453pp; English.  
XX  
CC The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridizes  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung cancer-  
CC associated polynucleotides and polypeptides are used for identifying a  
CC compound that modulates a lung cancer-associated polypeptide, for  
CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
CC cancer in a patient and for treating a mammal having lung cancer by  
CC administering a modulatory compound identified. The methods are useful  
CC for treating lung cancer, such as small cell lung cancer, non-small cell  
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
CC for diagnostic purposes and as targets for screening for therapeutic  
CC compounds that modulate lung cancer, such as antibodies. Sequences  
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the  
XX invention  
XX  
SQ Sequence 784 AA;

Query Match 84.9%; Score 73; DB 6; Length 784;  
Best Local Similarity 92.9%; Pred. No. 0.0064;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 14  
 DB 48 DWINWQHIDEKN 61

RESULT 14  
 ADA03363  
 ID ADA03363 standard; protein; 784 AA.

AC ADA03363;

DT 06-NOV-2003 (first entry)

DE Human vascular endothelial cadherin protein.

XX cytostratic; antidiabetic; ophthalmological; antiarthritic;

KW endostatin receptor; angiogenesis inhibitor; human;

KW vascular endothelial cadherin; angiogenesis; blood vessel; cancer;

KW diabetic retinopathy; rheumatoid arthritis; receptor.

OS Homo sapiens.

XX MO2003038096-A1.

XX 08-MAY-2003.

XX 28-OCT-2002; 2002WO-JP011143.

XX 30-OCT-2001; 2001JP-00333295.

XX (KUBO/) KUBOTA S.

XX (NEMO/) NEMOTO T.

XX Kubota S, Nemoto T;

XX WPI: 2003-430528/40.

XX N-PSDB; ADA03362.

XX Vascular endothelial cadherin is a receptor to endostatin for screening

XX endostatin agonists and antagonists as angiogenesis inhibitors and

XX remedies for cancer development and proliferation.

XX Claim 2; Page 44-47; 56pp; Japanese.

XX The invention relates to a DNA encoding an endostatin receptor protein

XX (human vascular endothelial (VE) cadherin (1)) or its partial peptides or

XX encoding a protein derived from (1) by addition, deletion or substitution

XX of one or more amino acid residues and having similar binding activity to

XX endostatin. The sequence can be used in the regulation of angiogenesis in

XX the treatment and prevention of diseases associated with the formation of

XX new blood vessels, including cancer, cancer proliferation, diabetic

XX retinopathy and rheumatoid arthritis. This sequence corresponds to the VE

XX cadherin protein which serves as an endostatin receptor protein.

XX Sequence 784 AA;

XX Query Match 84.9%; Score 73; DB 6; Length 784;

XX Best Local Similarity 92.9%; Pred. No. 0.0064;

XX Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 14

DB 48 DWINWQHIDEKN 61

RESULT 15

ABR49545

ID ABR49545 standard; protein; 77 AA.

AC ABR49545;

XX 05-FEB-2002 (first entry)

XX Listeria monocytogenes protein #2249.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

XX vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

XX MO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR001118.

XX 11-APR-2000; 2000FR-00004629.

XX (INSP) INST PASTER.

XX Buchrieser C, Frangoul L, Couve E, Rusniok C, Feihl H, Deboux P;

XX Dussurget O, Chetoui F, Nedjati H, Glaser P, Kunz F, Cosserat P;

XX Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

XX Dominguez-Bernal G, Garrido-Garcia P, Tleitez-Martinez A, Amend A;

XX Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;

XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

XX Madueno E, De Pablo B, Kaerst U, Entian K, Hauf J;

XX Rose M, Voess H;

XX WPI: 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

XX and prevention of Listeria and related bacterial infections, and related

XX polypeptides.

XX Claim 6; SEQ ID NO 2250; 192pp; French.

XX The present invention relates to the genome sequence of Listeria

XX monocytogenes EGD-e (see ABR49545). The genome sequence and fragments of

XX it are useful for selecting probes and primers for detecting genes in L.

XX monocytogenes and other genomes. The present invention is a protein

XX encoded by the genome sequence of the present invention. Proteins

XX expressed from the genome sequence are useful for raising specific

XX antibodies, identification of L. monocytogenes and related organisms, and

XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin

XX B12. The genome sequence and proteins encoded by it are also useful for

XX selecting compounds that regulate gene expression and cell replication

XX and modulate L. monocytogenes-related diseases. In addition, the genome

XX sequence and proteins encoded by it are useful in pharmaceutical and

XX vaccine compositions for the treatment or prevention of infections by L.

XX monocytogenes and related organisms. Note: The sequence data for this

XX patent did not form part of the printed specification, but was obtained

XX in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 77 AA;

XX Query Match 58.1%; Score 50; DB 5; Length 77;

XX Best Local Similarity 53.3%; Pred. No. 2.2;

XX Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEINWQHIDEKN 15

DB 62 DELWDAYLDEKNK 76

Search completed: July 29, 2004, 11:34:52

JOB time : 48.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:37:16 ; Search time 38.3333 Seconds  
(without alignments)  
122.745 Million cell updates/sec

Title: US-10-040-128-2

Perfect score: 92  
Sequence: 1 DWINQMHIDEKNE 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1291235 segs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	15	US-10-040-128-2	Sequence 2, Appl1
2	87	94.6	15	US-10-040-128-3	Sequence 3, Appl1
3	87	94.6	108	US-10-006-869-6	Sequence 6, Appl1
4	87	94.6	108	US-10-395-032-6	Sequence 6, Appl1
5	87	94.6	784	US-09-905-983-52	Sequence 52, Appl1
6	87	94.6	784	US-10-211-462-105	Sequence 105, Appl1
7	87	94.6	784	US-10-021-660-99	Sequence 99, Appl1
8	87	94.6	784	US-10-174-677-23	Sequence 23, Appl1
9	78	84.8	15	US-10-040-128-1	Sequence 1, Appl1
10	52	56.5	167	US-10-243-552-596	Sequence 596, App
11	52	56.5	772	US-09-978-295A-264	Sequence 264, App
12	52	56.5	772	US-09-978-687-264	Sequence 264, App
13	52	56.5	772	US-09-978-182A-264	Sequence 264, App
14	52	56.5	772	US-09-999-832A-264	Sequence 264, App
15	52	56.5	772	US-09-978-189-264	Sequence 264, App

16	52	56.5	772	10	US-09-978-608A-264	Sequence 264, App
17	52	56.5	772	10	US-09-978-585A-264	Sequence 264, App
18	52	56.5	772	10	US-09-978-191A-264	Sequence 264, App
19	52	56.5	772	10	US-09-978-403A-264	Sequence 264, App
20	52	56.5	772	10	US-09-978-564A-264	Sequence 264, App
21	52	56.5	772	10	US-09-999-833A-264	Sequence 264, App
22	52	56.5	772	10	US-09-981-915A-264	Sequence 264, App
23	52	56.5	772	10	US-09-978-824-264	Sequence 264, App
24	52	56.5	772	10	US-09-918-588A-264	Sequence 264, App
25	52	56.5	772	10	US-09-978-423A-264	Sequence 264, App
26	52	56.5	772	10	US-09-978-187B-264	Sequence 264, App
27	52	56.5	772	10	US-09-999-830A-264	Sequence 264, App
28	52	56.5	772	10	US-09-978-757A-264	Sequence 264, App
29	52	56.5	772	10	US-09-978-643A-264	Sequence 264, App
30	52	56.5	772	10	US-09-978-375A-264	Sequence 264, App
31	52	56.5	772	10	US-09-978-188A-264	Sequence 264, App
32	52	56.5	772	10	US-09-978-194A-264	Sequence 264, App
33	52	56.5	772	10	US-09-978-681A-264	Sequence 264, App
34	52	56.5	772	10	US-09-978-828A-264	Sequence 264, App
35	52	56.5	772	10	US-09-978-299A-264	Sequence 264, App
36	52	56.5	772	10	US-09-978-665A-264	Sequence 264, App
37	52	56.5	772	10	US-09-978-802A-264	Sequence 264, App
38	52	56.5	772	10	US-10-164-749A-264	Sequence 264, App
39	52	56.5	772	10	US-10-206-915-106	Sequence 106, App
40	52	56.5	772	12	US-10-199-670-106	Sequence 106, App
41	52	56.5	772	12	US-10-201-858-106	Sequence 106, App
42	52	56.5	772	12	US-09-999-831A-264	Sequence 264, App
43	52	56.5	772	12		
44	52	56.5	772	12		
45	52	56.5	772	12		

## ALIGNMENTS

RESULT 1  
US-10-040-128-2  
; Sequence 2, Application US/10040128  
; Publication No. US20020160003A1  
; GENERAL INFORMATION:  
; APPLICANT: Liao, Fang  
; APPLICANT: Hicklin, Daniel J.  
; TITLE OF INVENTION: Antibody Antagonists of VE-Cadherin Without Adverse Effects on  
; TITLE OF INVENTION: Vascular Permeability  
; FILE REFERENCE: 11245/46976  
; CURRENT APPLICATION NUMBER: US/10/040,128  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-10-040-128-2

Query Match  
Best Local Similarity 100.0%; Score 92; DB 13; Length 15;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINQMHIDEKNE 15  
|||  
Db 1 DWINQMHIDEKNE 15

RESULT 2  
US-10-040-128-3  
; Sequence 3, Application US/10040128  
; Publication No. US20020160003A1  
; GENERAL INFORMATION:  
; APPLICANT: Liao, Fang

```
; APPLICANT: Hicklin, Daniel J.
; APPLICANT: Bohlen, Peter
; TITLE OF INVENTION: Anticbody Antagonists of VE-Cadherin Without Adverse Effects on
; FILE REFERENCE: 11245/46976
; CURRENT APPLICATION NUMBER: US/10/040,128
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-040-128-3

Query Match          94.6%; Score 87; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  DWIMNQMHIDEKN 14
      |||||
Db      1  DWIMNQMHIDEKN 14

RESULT 3
US-10-006-869-6
; Sequence 6, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-869-6

Query Match          94.6%; Score 87; DB 14; Length 108;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  DWIMNQMHIDEKN 14
      |||||
Db      5  DWIMNQMHIDEKN 18

RESULT 4
US-10-395-032-6
; Sequence 6, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-395-032-6

Query Match          94.6%; Score 87; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  DWIMNQMHIDEKN 14
      |||||
Db      5  DWIMNQMHIDEKN 18

RESULT 5
US-09-905-983-52
; Sequence 52, Application US/09905983
; Patent No. US20020045591A1
; GENERAL INFORMATION:
; APPLICANT: Geisler, Benjamin
; APPLICANT: Ben-Ze'ev, Avri
; APPLICANT: Sadot, Eliaht
; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER
; FILE REFERENCE: 01/22326
; CURRENT APPLICATION NUMBER: US/09/905,983
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-983-52

Query Match          94.6%; Score 87; DB 9; Length 784;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  DWIMNQMHIDEKN 14
      |||||
Db      48 DWIMNQMHIDEKN 61

RESULT 6
US-10-211-462-105
; Sequence 105, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-105
```



Query Match 94.6%; Score 87; DB 12; Length 784;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINQMHIDEKN 14  
|||||  
Db 48 DWINQMHIDEKN 61

## RESULT 7

US-10-021-660-99  
; Sequence 99, Application US/10021660  
; Publication No. US20030152926A1  
; GENERAL INFORMATION:  
; APPLICANT: Murray, Richard  
; APPLICANT: Gurney, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: ROS Biotechnology, Inc.  
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,  
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis  
; FILE REFERENCE: 018501-00071005  
; CURRENT APPLICATION NUMBER: US/10/021,660  
; CURRENT FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: US/09/784,356  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: US 09/637,977  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 99  
; LENGTH: 784  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-021-660-99

Query Match 94.6%; Score 87; DB 14; Length 784;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINQMHIDEKN 14  
|||||  
Db 48 DWINQMHIDEKN 61

## RESULT 8

US-10-174-677-23  
; Sequence 23, Application US/10174677  
; Publication No. US20030190704A1  
; GENERAL INFORMATION:  
; APPLICANT: Xie, Ting  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR  
; FILE REFERENCE: 40716(IP-012)  
; CURRENT APPLICATION NUMBER: US/10/174,677  
; CURRENT FILING DATE: 2002-06-19  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 784  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-174-677-23

Query Match 94.6%; Score 87; DB 14; Length 784;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINQMHIDEKN 14  
|||||  
Db 48 DWINQMHIDEKN 61

## RESULT 9

US-10-040-128-1  
; Sequence 1, Application US/10040128  
; Publication No. US20020160003A1  
; GENERAL INFORMATION:  
; APPLICANT: Liao, Fang  
; APPLICANT: Hicklin, Daniel J.  
; APPLICANT: Bohlen, Peter  
; TITLE OF INVENTION: Antibody Antagonists of VE-Cadherin Without Adverse Effects on  
; FILE REFERENCE: 11245/46976  
; CURRENT APPLICATION NUMBER: US/10/040,128  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 1  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-10-040-128-1

Query Match 84.8%; Score 78; DB 13; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00012;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DWINQMHIDEKNE 15  
|||||  
Db 1 DWINQMHIDEKNE 15

RESULT 10  
US-10-243-552-596  
; Sequence 596, Application US/10243552  
; Publication No. US20030224379A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Weng, Gezhi  
; APPLICANT: Ma, Yundong  
; TITLE OF INVENTION: Novel Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 807A  
; CURRENT APPLICATION NUMBER: US/10/243,552  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 60/322,511  
; PRIOR FILING DATE: 2001-09-13  
; PRIOR APPLICATION NUMBER: PCT/US00/35017  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: PCT/US01/02623  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: US 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: PCT/US01/03800  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: US 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US01/04927  
; PRIOR FILING DATE: 2001-02-26  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 998  
; SOFTWARE: pc\_fl\_genes Version 5.0  
; SEQ ID NO 596  
; LENGTH: 167  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-243-552-596

Query Match	56.5%	Score 52;	DB 12;	Length 167;
Best Local Similarity	53.8%	Pred. No. 7.2;		
Matches	7;	Conservative	2;	Mismatches 4; Indels 0; Gaps 0

QY	2	W	I	W	N	Q	M	H	I	D	E	K	N	14
			:				:							
Db	45	W	V	W	N	Q	F	F	V	P	E	E	M	57

RESULT 11  
US-09-978

```

1 / Sequence 264, Application US/09978295A
2 / Patent No. US20020156006A1
3 /
4 / GENERAL INFORMATION:
5 / APPLICANT: Ashkenazi, Avi
6 / APPLICANT: Baker Kevin P.
7 / APPLICANT: Borstein, David
8 / APPLICANT: Desnoyers, Luc
9 / APPLICANT: Eaton, Dan
10 / APPLICANT: Ferrara, Napoleon
11 / APPLICANT: Filvaroff, Ellen
12 / APPLICANT: Fong, Sherman
13 / APPLICANT: Gao, Wei-Qiang
14 / APPLICANT: Gerber, Hanspeter
15 / APPLICANT: Gertsen, Mary E.
16 / APPLICANT: Goddard, Audrey
17 / APPLICANT: Godowski, Paul J.
18 / APPLICANT: Grimaldi, J. Christopher
19 / APPLICANT: Gurney, Austin L.
20 / APPLICANT: Hillan, Kenneth J.
21 / APPLICANT: Kljavin, Ivar J.
22 / APPLICANT: Kuo, Sophia S.
23 / APPLICANT: Napier, Mary A.
24 / APPLICANT: Pan, James;
25 / APPLICANT: Paoni, Nicholas F.
26 / APPLICANT: Roy, Margaret Ann
27 / APPLICANT: Shelton, David L.
28 / APPLICANT: Stewart, Timothy A.
29 / APPLICANT: Tumas, Daniel
30 / APPLICANT: Williams, P. Mickey
31 /
32 / TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
33 /
34 / FILE REFERENCE: P2630P1C11
35 /
36 / CURRENT APPLICATION NUMBER: US/09/978,295A
37 /
38 / CURRENT FILING DATE: 2001-10-15
39 /
40 / PRIOR APPLICATION NUMBER: 09/918585
41 /
42 / PRIOR FILING DATE: 2001-07-30
43 /
44 / PRIOR APPLICATION NUMBER: 60/062250
45 /
46 / PRIOR FILING DATE: 1997-10-17
47 /
48 / PRIOR APPLICATION NUMBER: 60/064249
49 /
50 / PRIOR FILING DATE: 1997-11-03
51 /
52 / PRIOR APPLICATION NUMBER: 60/065311
53 /
54 / PRIOR FILING DATE: 1997-11-13
55 /
56 / PRIOR APPLICATION NUMBER: 60/066364
57 /
58 / PRIOR FILING DATE: 1997-11-21
59 /
60 / PRIOR APPLICATION NUMBER: 60/077450
61 /
62 / PRIOR FILING DATE: 1998-03-10
63 /
64 / PRIOR APPLICATION NUMBER: 60/077632
65 /
66 / PRIOR FILING DATE: 1998-03-11
67 /
68 / PRIOR APPLICATION NUMBER: 60/077641
69 /
70 / PRIOR FILING DATE: 1998-03-11
71 /
72 / PRIOR APPLICATION NUMBER: 60/077649
73 /
74 / PRIOR FILING DATE: 1998-03-11
75 /
76 / PRIOR APPLICATION NUMBER: 60/077791
77 /
78 / PRIOR FILING DATE: 1998-03-12
79 /
80 / PRIOR APPLICATION NUMBER: 60/078004
81 /
82 / PRIOR FILING DATE: 1998-03-13
83 /
84 / PRIOR APPLICATION NUMBER: 60/078886
85 /
86 / PRIOR FILING DATE: 1998-03-20
87 /
88 / PRIOR APPLICATION NUMBER: 60/078936
89 /

```

1	PRIOR FILING DATE: 1998-03-20	
2	PRIOR APPLICATION NUMBER: 60/078910	
3	PRIOR FILING DATE: 1998-03-20	
4	PRIOR APPLICATION NUMBER: 60/078939	
5	PRIOR FILING DATE: 1998-03-20	
6	PRIOR APPLICATION NUMBER: 60/079294	
7	PRIOR FILING DATE: 1998-03-25	
8	PRIOR APPLICATION NUMBER: 60/079656	
9	PRIOR FILING DATE: 1998-03-26	
10	PRIOR APPLICATION NUMBER: 60/079664	
11	PRIOR FILING DATE: 1998-03-27	
12	PRIOR APPLICATION NUMBER: 60/079689	
13	PRIOR FILING DATE: 1998-03-27	
14	PRIOR APPLICATION NUMBER: 60/079728	
15	PRIOR FILING DATE: 1998-03-27	
16	PRIOR APPLICATION NUMBER: 60/079786	
17	PRIOR FILING DATE: 1998-03-27	
18	PRIOR APPLICATION NUMBER: 60/079920	
19	PRIOR FILING DATE: 1998-03-30	
20	PRIOR APPLICATION NUMBER: 60/079923	
21	PRIOR FILING DATE: 1998-03-30	
22	PRIOR APPLICATION NUMBER: 60/080105	
23	PRIOR FILING DATE: 1998-03-31	
24	PRIOR APPLICATION NUMBER: 60/080107	
25	PRIOR FILING DATE: 1998-03-31	
26	PRIOR APPLICATION NUMBER: 60/080165	
27	PRIOR FILING DATE: 1998-03-31	
28	PRIOR APPLICATION NUMBER: 60/080194	
29	PRIOR FILING DATE: 1998-03-31	
30	PRIOR APPLICATION NUMBER: 60/080327	
31	PRIOR FILING DATE: 1998-04-01	
32	PRIOR APPLICATION NUMBER: 60/080328	
33	PRIOR FILING DATE: 1998-04-01	
34	PRIOR APPLICATION NUMBER: 60/080333	
35	PRIOR FILING DATE: 1998-04-01	
36	PRIOR APPLICATION NUMBER: 60/080334	
37	PRIOR FILING DATE: 1998-04-01	
38	PRIOR APPLICATION NUMBER: 60/081070	
39	PRIOR FILING DATE: 1998-04-08	
40	PRIOR APPLICATION NUMBER: 60/081049	
41	PRIOR FILING DATE: 1998-04-08	
42	PRIOR APPLICATION NUMBER: 60/081071	
43	PRIOR FILING DATE: 1998-04-09	
44	PRIOR APPLICATION NUMBER: 60/081195	
45	PRIOR FILING DATE: 1998-04-08	
46	PRIOR APPLICATION NUMBER: 60/081203	
47	PRIOR FILING DATE: 1998-04-09	
48	PRIOR APPLICATION NUMBER: 60/081229	
49	PRIOR FILING DATE: 1998-04-09	
50	PRIOR APPLICATION NUMBER: 60/081555	
51	PRIOR FILING DATE: 1998-04-15	
52	PRIOR APPLICATION NUMBER: 60/081817	
53	PRIOR FILING DATE: 1998-04-15	
54	PRIOR APPLICATION NUMBER: 60/081819	
55	PRIOR FILING DATE: 1998-04-15	
56	PRIOR APPLICATION NUMBER: 60/081952	
57	PRIOR FILING DATE: 1998-04-15	
58	PRIOR APPLICATION NUMBER: 60/081952	
59	PRIOR FILING DATE: 1998-04-15	
60	PRIOR APPLICATION NUMBER: 60/081838	
61	PRIOR FILING DATE: 1998-04-15	
62	PRIOR APPLICATION NUMBER: 60/082568	
63	PRIOR FILING DATE: 1998-04-21	
64	PRIOR APPLICATION NUMBER: 60/082569	
65	PRIOR FILING DATE: 1998-04-21	
66	PRIOR APPLICATION NUMBER: 60/082704	
67	PRIOR FILING DATE: 1998-04-22	
68	PRIOR APPLICATION NUMBER: 60/082804	
69	PRIOR FILING DATE: 1998-04-22	
70	PRIOR APPLICATION NUMBER: 60/082700	
71	PRIOR FILING DATE: 1998-04-22	
72	PRIOR APPLICATION NUMBER: 60/082797	
73	PRIOR FILING DATE: 1998-04-22	
74	PRIOR APPLICATION NUMBER: 60/082797	

;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/08579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/08573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.5%; Score 52; DB 9; Length 772;  
Best Local Similarity 53.8%; Pred. No. 31;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
2 WINNOMHIDEKN 14

Db 45 WYWNQFVPEEMN 57

RESULT 12  
US-09-978-697-264  
; Sequence 264, Application US/09978697  
; Patent No. US2002016284A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Guiney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC27  
; CURRENT APPLICATION NUMBER: US/09/978,697  
;; PRIOR APPLICATION NUMBER: 2001-10-16  
;; PRIOR FILING DATE: 2001-07-30  
;; PRIOR APPLICATION NUMBER: 60/062250  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064249  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/065311  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/077450  
;; PRIOR FILING DATE: 1998-03-10  
;; PRIOR APPLICATION NUMBER: 60/077632  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077641  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077649  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078004  
;; PRIOR FILING DATE: 1998-03-13  
;; PRIOR APPLICATION NUMBER: 60/078886  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078936  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392

PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.5%; Score 52; DB 9; Length 772;  
Best Local Similarity 53.8%; Pred. No. 31;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 WTMNQHHIDEKN 14  
DB 45 WYWNQFVPEKN 57

RESULT 13  
US-09-978-192A-264  
; Sequence 264, Application US/09978192A

```

1 Patent No. US20020177553A1
2 GENERAL INFORMATION:
3 APPLICANT: Ashkenazi, Avi
4 APPLICANT: Baker Kevin P.
5 APPLICANT: Botstein, David
6 APPLICANT: Desnovers, Luc
7 APPLICANT: Eaton, Dan
8 APPLICANT: Ferrara, Napoleon
9 APPLICANT: Filvaroff, Ellen
10 APPLICANT: Fong, Sherman
11 APPLICANT: Gao, Wei-Qiang
12 APPLICANT: Geber, Hanspeter
13 APPLICANT: Gerltsen, Mary E.
14 APPLICANT: Goddard, Audrey
15 APPLICANT: Godowski, Paul J.
16 APPLICANT: Grimaldi, J. Christopher
17 APPLICANT: Gurney, Austin L.
18 APPLICANT: Hillan, Kenneth J.
19 APPLICANT: Kljavin, Ivar J.
20 APPLICANT: Kuo, Sophia S.
21 APPLICANT: Napier, Mary A.
22 APPLICANT: Paoni, James,
23 APPLICANT: Paoni, Nicholas F.
24 APPLICANT: Roy, Margaret Ann
25 APPLICANT: Shelton, David L.
26 APPLICANT: Stewart, Timothy A.
27 APPLICANT: Tumas, Daniel
28 APPLICANT: Williams, P. Mickey
29 APPLICANT: Wood, William I.
30 TITLE OF INVENTION: Acids and Transmembrane Polypeptides and Nucleic
31 FILE REFERENCE: P2630PIC9
32 CURRENT APPLICATION NUMBER: US/09/979,192A
33 CURRENT FILING DATE: 2001-10-15
34 PRIOR APPLICATION NUMBER: 09/918585
35 PRIOR FILING DATE: 2001-07-30
36 PRIOR APPLICATION NUMBER: 60/062250
37 PRIOR FILING DATE: 1997-10-17
38 PRIOR APPLICATION NUMBER: 60/064249
39 PRIOR FILING DATE: 1997-11-03
40 PRIOR APPLICATION NUMBER: 60/065311
41 PRIOR FILING DATE: 1997-11-13
42 PRIOR APPLICATION NUMBER: 60/066364
43 PRIOR FILING DATE: 1997-11-21
44 PRIOR APPLICATION NUMBER: 60/077450
45 PRIOR FILING DATE: 1998-03-10
46 PRIOR APPLICATION NUMBER: 60/077632
47 PRIOR FILING DATE: 1998-03-11
48 PRIOR APPLICATION NUMBER: 60/077641
49 PRIOR FILING DATE: 1998-03-11
50 PRIOR APPLICATION NUMBER: 60/077649
51 PRIOR FILING DATE: 1998-03-11
52 PRIOR APPLICATION NUMBER: 60/077791
53 PRIOR FILING DATE: 1998-03-12
54 PRIOR APPLICATION NUMBER: 60/078004
55 PRIOR FILING DATE: 1998-03-13
56 PRIOR APPLICATION NUMBER: 60/078886
57 PRIOR FILING DATE: 1998-03-20
58 PRIOR APPLICATION NUMBER: 60/078936
59 PRIOR FILING DATE: 1998-03-20
60 PRIOR APPLICATION NUMBER: 60/078910
61 PRIOR FILING DATE: 1998-03-20
62 PRIOR APPLICATION NUMBER: 60/078939
63 PRIOR FILING DATE: 1998-03-20
64 PRIOR APPLICATION NUMBER: 60/079294
65 PRIOR FILING DATE: 1998-03-25
66 PRIOR APPLICATION NUMBER: 60/079656
67 PRIOR FILING DATE: 1998-03-26
68 PRIOR APPLICATION NUMBER: 60/079664
69 PRIOR FILING DATE: 1998-03-27
70 PRIOR APPLICATION NUMBER: 60/079669
71 PRIOR FILING DATE: 1998-03-27
72 PRIOR APPLICATION NUMBER: 60/079663

```

PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079572
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079786
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079922
PRIOR FILING DATE:	1998-03-30
PRIOR APPLICATION NUMBER:	60/079922
PRIOR FILING DATE:	1998-03-30
PRIOR APPLICATION NUMBER:	60/080165
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080199
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080332
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080332
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080333
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080344
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/081070
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081049
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081071
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081195
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081203
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081229
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081955
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081817
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081819
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081952
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081838
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/082568
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/082569
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/082704
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/082804
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/082700
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/082797
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/082796
PRIOR FILING DATE:	1998-04-23
PRIOR APPLICATION NUMBER:	60/083336
PRIOR FILING DATE:	1998-04-27
PRIOR APPLICATION NUMBER:	60/083322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/083392
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083495
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083496
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083499
PRIOR FILING DATE:	1998-04-29

```

; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR APPLICATION NUMBER: 60/085697

Query Match          56.5%; Score 52; DB 9; Length 772;
Best Local Similarity 53.8%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 WTWNQMHIDEKXN 14
DB      45 WYWNQFVPEEKXN 57

RESULT 14
US-09-999-832A-264
; Sequence 264, Application US/0999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
```

```

; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Rong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C63
; CURRENT APPLICATION NUMBER: US/09/999,832A
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
```

PRIOR APPLICATION NUMBER: 60/079923  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/080105  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080107  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559

PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084441  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084639  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084640  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084598  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084643  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085582  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085700  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085689  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085580  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085573  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.5%; Score 52; DB 9; Length 772;  
 Best Local Similarity 53.8%; Pred. No. 31;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 WMMNMHIDEKN 14  
 DB 45 WMMNFFVPEEKN 57

RESULT 15  
 US-09-978-189-264  
 Sequence 264; Application US/09978189  
 Publication No. US2003004102A1  
 GENERAL INFORMATION:  
 APPLICANT: Ashkenazi, Avi  
 APPLICANT: Baker Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Deenoyers, Luc  
 APPLICANT: Eaton, Dan  
 APPLICANT: Ferrara, Napoleon  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerber, Hanspeter  
 APPLICANT: Gottlieb, Mary E.  
 APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: KJavian, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James J.  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Thomas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C7  
CURRENT APPLICATION NUMBER: US/09/978,189  
PRIOR FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165

PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05



/ PRIOR APPLICATION NUMBER: 60/084414  
 / PRIOR FILING DATE: 1998-05-06  
 / PRIOR APPLICATION NUMBER: 60/084441  
 / PRIOR FILING DATE: 1998-05-06  
 / PRIOR APPLICATION NUMBER: 60/084637  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/084639  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/084640  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/084598  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/084600  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/084627  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/084643  
 / PRIOR FILING DATE: 1998-05-07  
 / PRIOR APPLICATION NUMBER: 60/085339  
 / PRIOR FILING DATE: 1998-05-13  
 / PRIOR APPLICATION NUMBER: 60/085338  
 / PRIOR FILING DATE: 1998-05-13  
 / PRIOR APPLICATION NUMBER: 60/085323  
 / PRIOR FILING DATE: 1998-05-13  
 / PRIOR APPLICATION NUMBER: 60/085582  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085700  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085689  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085579  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085580  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085573  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085704  
 / PRIOR FILING DATE: 1998-05-15  
 / PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.5%; Score 52; DB 10; Length 772;  
 Best Local Similarity 53.8%; Pred. No. 31;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 WTWNMHIDEKN 14  
 Db 45 WVMNQFVDEKN 57

Search completed: July 29, 2004, 11:50:32  
 Job time : 38.333 secs

**This Page Blank (uspic)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:32:25 ; Search time 13 Seconds

(Without alignments)  
59.568 Million cell updates/sec

Title: US-10-040-128-2

Sequence: 1 DWIMNQHIDEKNE 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/6C\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/6D\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	94.6	108	4	US-09-187-859-6
2	87	94.6	108	4	US-09-839-542B-6
3	87	94.6	108	4	US-09-535-852-6
4	87	94.6	780	1	US-08-188-228-50
5	87	94.6	780	1	US-08-332-643-44
6	87	94.6	780	1	US-08-332-638-50
7	49	53.3	110	4	US-09-187-859-13
8	49	53.3	110	4	US-09-839-542B-13
9	49	53.3	532	1	US-08-188-228-44
10	49	53.3	532	1	US-08-332-638-44
11	49	53.3	793	1	US-08-188-228-54
12	49	53.3	793	1	US-08-332-643-48
13	49	53.3	793	1	US-08-332-638-54
14	49	53.3	799	1	US-08-188-228-42
15	49	53.3	799	1	US-08-332-638-42
16	49	53.3	799	1	US-09-187-859-44
17	45	48.9	106	4	US-09-187-859-45
18	45	48.9	106	4	US-09-305-927-1
19	45	48.9	106	4	US-09-305-927-2
20	45	48.9	106	4	US-09-305-927-3
21	45	48.9	106	4	US-09-073-040-12
22	45	48.9	106	4	US-09-073-040-13
23	45	48.9	106	4	US-09-839-542B-44
24	45	48.9	106	4	US-09-839-542B-45
25	45	48.9	106	4	US-09-264-516A-1
26	45	48.9	106	4	US-09-264-516A-2
27	45	48.9	106	4	US-09-264-516A-3

28	45	48.9	106	4	US-09-264-516A-3	Sequence 3, Appli
29	45	48.9	106	4	US-09-234-395-4	Sequence 4, Appli
30	45	48.9	106	4	US-09-234-395-5	Sequence 5, Appli
31	45	48.9	106	4	US-09-305-928-4	Sequence 4, Appli
32	45	48.9	106	4	US-09-305-928-5	Sequence 5, Appli
33	45	48.9	110	4	US-09-187-859-4	Sequence 4, Appli
34	45	48.9	110	4	US-09-187-859-22	Sequence 22, Appli
35	45	48.9	110	4	US-09-839-542B-4	Sequence 4, Appli
36	45	48.9	110	4	US-09-839-542B-22	Sequence 22, Appli
37	45	48.9	110	4	US-09-535-852-4	Sequence 4, Appli
38	45	48.9	110	4	US-09-535-852-22	Sequence 22, Appli
39	45	48.9	615	2	US-08-738-349-12	Sequence 12, Appli
40	45	48.9	693	2	US-08-738-349-6	Sequence 6, Appli
41	45	48.9	796	1	US-08-188-228-58	Sequence 58, Appli
42	45	48.9	796	1	US-08-332-643-52	Sequence 52, Appli
43	45	48.9	796	1	US-08-332-638-58	Sequence 58, Appli
44	45	48.9	796	2	US-08-738-349-2	Sequence 2, Appli
45	45	48.9	796	2	US-08-738-349-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-09-187-859-6  
; Sequence 6, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187, 859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-187-859-6

Query Match 94.6% Score 87; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMNQHIDEKNE 14  
DB 5 DWIMNQHIDEKNE 18

RESULT 2  
US-09-839-542B-6  
; Sequence 6, Application US/09839542B  
; Patent No. 656996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839, 542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-839-542B-6

Query Match 94.6%; Score 87; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 4,4e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIWMQMHIDEKN 14  
|||||  
DB 5 DWIWMQMHIDEKN 18

## RESULT 3

US-09-535-852-6  
; Sequence 6, Application US/09535852  
; Patent No. 6638911  
; GENERAL INFORMATION:  
; APPLICANT: Blachuk, Orest W.  
; APPLICANT: Symonds, James M.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING  
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C6  
; CURRENT APPLICATION NUMBER: US/09/535,852  
; CURRENT FILING DATE: 2001-05-21  
; NUMBER OF SEQ ID NOS: 2009  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-535-852-6

Query Match 94.6%; Score 87; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 4,4e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIWMQMHIDEKN 14  
|||||  
DB 5 DWIWMQMHIDEKN 18

## RESULT 4

US-08-188-228-50  
; Sequence 50, Application US/08188228  
; Patent No. 5597725  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray &  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/188,228  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/049,460  
; FILING DATE: 19 APR 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,643  
; FILING DATE: 17 APR 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5597725and, Greta E.  
; REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31340  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELE: 25-3856  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-188-228-50

Query Match 94.6%; Score 87; DB 1; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3,8e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIWMQMHIDEKN 14  
|||||  
DB 44 DWIWMQMHIDEKN 57

## RESULT 5

US-08-332-643-44  
; Sequence 44, Application US/08332643  
; Patent No. 5639634  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray &  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,643  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/872,643  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5639634and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELE: 25-3856  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-332-643-44

Query Match 94.6%; Score 87; DB 1; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3,8e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIWMQMHIDEKN 14  
|||||

Db 44 DWINQMHIDEKN 57

RESULT 6  
US-08-332-638-50  
; Sequence 50, Application US/08332638  
; Patent No. 5646250  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,638  
; FILING DATE: 01-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,643  
; FILING DATE: 17 APR 1992  
; APPLICATION NUMBER: US/08/049,460  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5646250and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31340  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-332-638-50

Query Match 94.6%; Score 87; DB 1; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3.8e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DWINQMHIDEKN 14  
Db 44 DWINQMHIDEKN 57  
RESULT 7  
US-09-187-859-13  
; Sequence 13, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 13

; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-187-859-13

Query Match 53.3%; Score 49; DB 4; Length 110;  
Best Local Similarity 63.6%; Pred. No. 2.2;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WWINQMHIDE 12  
Db 6 WWINQMFVLEE 16

RESULT 8  
US-09-839-542B-13  
; Sequence 13, Application US/09839542B  
; Patent No. 656996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-839-542B-13

Query Match 53.3%; Score 49; DB 4; Length 110;  
Best Local Similarity 63.6%; Pred. No. 2.2;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WWINQMHIDE 12  
Db 6 WWINQMFVLEE 16

RESULT 9  
US-09-535-852-13  
; Sequence 13, Application US/09535852  
; Patent No. 6638911  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James M.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING  
; FILE REFERENCE: 100086.407C6  
; CURRENT APPLICATION NUMBER: US/09/535,852  
; CURRENT FILING DATE: 2001-05-21  
; NUMBER OF SEQ ID NOS: 2009  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-535-852-13

Query Match 53.3%; Score 49; DB 4; Length 110;  
Best Local Similarity 63.6%; Pred. No. 2.2;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WWINQMHIDE 12  
Db 6 WWINQMFVLEE 16

RESULT 10  
US-08-188-228-44  
Sequence 44, Application US/08188228  
Patent No. 5597725  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,228  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,460  
FILING DATE: 19 APR 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5597725and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-228-44

Query Match 53.3%; Score 49; DB 1; Length 532;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 MINOMHIDE 12  
|:|||||:|  
Db 63 WWMQMFVLEE 73

RESULT 11  
US-08-332-638-44  
Sequence 44, Application US/08332638  
Patent No. 5646250  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,638  
FILING DATE: 01-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
APPLICATION NUMBER: US/08/049,460  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5646250and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-332-638-44

Query Match 53.3%; Score 49; DB 1; Length 532;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 MINOMHIDE 12  
|:|||||:|  
Db 63 WWMQMFVLEE 73

RESULT 12  
US-08-188-228-54  
Sequence 54, Application US/08188228  
Patent No. 5597725  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,228  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,460  
FILING DATE: 19 APR 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5597725and, Greta E.  
REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-228-54

Query Match 53.3%; Score 49; DB 1; Length 793;  
Best Local Similarity 63.6%; Pred. No. 19;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12  
DB 56 WVNOMFVLEE 66

RESULT 13  
US-08-332-643-48  
Sequence 48, Application US/08332643  
Patent No. 5639634  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,643  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/872,643  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5639634and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30795  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-332-643-48

Query Match 53.3%; Score 49; DB 1; Length 793;  
Best Local Similarity 63.6%; Pred. No. 19;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 WINNOMHIDE 12  
DB 56 WVNOMFVLEE 66

DB 56 WVNOMFVLEE 66

RESULT 14  
US-08-332-638-54  
Sequence 54, Application US/08332638  
Patent No. 5646250  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,638  
FILING DATE: 01-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
APPLICATION NUMBER: US/08/049,460  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5646250and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-332-638-54

Query Match 53.3%; Score 49; DB 1; Length 793;  
Best Local Similarity 63.6%; Pred. No. 19;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12  
DB 56 WVNOMFVLEE 66

RESULT 15  
US-08-188-228-42  
Sequence 42, Application US/08188228  
Patent No. 5597725  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA

ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,228  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,460  
FILING DATE: 19 APR 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 559725and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 799 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-228-42

Query Match 53.3%; Score 49; DB 1; Length 799;  
Best Local Similarity 63.6%; Pred. NO. 19;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WNNQMHIDE 12  
|:||||:|  
Db 63 WNNQMFVLEE 73

Search completed: July 29, 2004, 11:38:39  
Job time : 13 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:30:55 ; Search time 31 Seconds

(Without alignments)  
152.670 Million cell updates/sec

Title: US-10-040-128-2  
Perfect score: 92  
Sequence: 1 DWINQMHIDEKNE 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP viirus:\*  
16: SP bacteriap:\*  
17: SP archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	94.6	782	6 Q8WNW5	Q8WNW5 sus scrofa
2	64	63.6	773	13 Q8AYD0	Q8AYD0 gallus galli
3	50	53.3	524	5 O01321	O01321 caenorhabdi
4	49	53.3	247	11 Q8C9N7	Q8C9N7 mus musculu
5	49	53.3	716	11 Q8C449	Q8C449 mus musculu
6	49	53.3	716	11 Q8C375	Q8C375 mus musculu
7	49	53.3	754	11 Q8BRK4	Q8BRK4 mus musculu
8	47	51.1	333	2 Q9L3G9	Q9L3G9 erwinia rna
9	46	50.0	313	16 Q88HL3	Q88HL3 pseudomonas
10	46	50.0	839	5 Q9N2R2	Q9N2R2 strongyloce
11	46	50.0	1497	16 Q87G80	Q87G80 vibrio para
12	45.5	49.5	143	5 Q9Y168	Q9Y168 drosophila
13	45	48.9	176	10 Q8RYE1	Q8RYE1 arena strig
14	45	48.9	241	4 Q8ET58	Q8ET58 homo sapien
15	45	48.9	241	5 Q8IK72	Q8IK72 plasmodium
16	45	48.9	310	10 Q8W5G3	Q8W5G3 oryza sativ

17	45	48.9	350	3 Q00870	Q00870 nectria hae
18	45	48.9	391	16 Q31773	Q31773 bacillus su
19	45	48.9	493	4 Q96L07	Q96L07 homo sapien
20	45	48.9	522	4 Q96DN1	Q96DN1 homo sapien
21	45	48.9	522	4 Q86UB6	Q86UB6 homo sapien
22	45	48.9	522	11 Q8X377	Q8X377 mus musculu
23	45	48.9	781	4 Q86UP1	Q86UP1 homo sapien
24	45	48.9	794	13 Q93264	Q93264 xenopus lae
25	45	48.9	796	4 Q96CZ9	Q96CZ9 homo sapien
26	45	48.9	796	11 Q8C7Q6	Q8C7Q6 mus musculu
27	45	48.9	819	4 Q86UP0	Q86UP0 homo sapien
28	45	48.9	1356	5 Q8ICG0	Q8ICG0 plasmodium
29	44.5	48.4	963	10 Q9STU7	Q9STU7 arabidopsis
30	44	47.8	76	11 Q8BLT5	Q8BLT5 mus musculu
31	44	47.8	163	16 Q99S13	Q99S13 staphylococ
32	44	47.8	164	2 Q9ZAH7	Q9ZAH7 staphylococ
33	44	47.8	219	16 Q8ECK3	Q8ECK3 shewanella
34	44	47.8	337	11 Q8BLB5	Q8BLB5 mus musculu
35	44	47.8	340	11 Q8BLT4	Q8BLT4 mus musculu
36	44	47.8	500	16 Q99T97	Q99T97 staphylococ
37	44	47.8	500	16 Q8NW11	Q8NW11 staphylococ
38	44	47.8	551	13 Q8AWM2	Q8AWM2 gallus galli
39	44	47.8	558	5 Q20375	Q20375 caenorhabdi
40	44	47.8	586	5 Q95ZT7	Q95ZT7 caenorhabdi
41	44	47.8	630	4 Q81Y78	Q81Y78 homo sapien
42	44	47.8	785	11 Q8BM92	Q8BM92 mus musculu
43	44	47.8	790	4 Q8N522	Q8N522 homo sapien
44	44	47.8	790	13 Q91838	Q91838 xenopus lae
45	44	47.8	794	4 Q86UD2	Q86UD2 homo sapien

## ALIGNMENTS

## RESULT 1

Q8WNW5 ID Q8WNW5 PRELIMINARY; PRT; 782 AA.  
AC Q8WNW5;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE VB cadherin.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_Taxid:9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hirano M., Hirano K., Nishimura J., Kanaide H.;  
RT "Transcriptional up-regulation of p27Kip1 during contact-induced  
RT growth arrest in the endothelial cells.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=99119339; PubMed=9918866;  
RA Hirano M., Niiro N., Hirano K., Nishimura J., Harshe D.J.,  
RA Kanaide H.;  
RT "Expression, subcellular localization and cloning of the 130 kDa  
RT regulatory subunit of myosin phosphatase in porcine aortic endothelial  
RL cells.";  
RL Biochem. Biophys. Res. Commun. 254:490-496(1999).  
CC -I- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
CC SORTING OF HETEROGENEOUS CELL TYPES (BY SIMILARITY).  
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
DR EMBL; AB046120; BAB82983.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005194; F:cell adhesion molecule activity; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.

```

DR Interpro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN-.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS02068; CADHERIN_2; 5.
DR KEGG; Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 782 AA; 87536 MW; 6E1D12D017CB1083 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMNOMHIDEKN 14
Db 45 DWIMNOMHIDEKN 58

RESULT 2
Q8AYDO PRELIMINARY; PRT; 773 AA.
ID 08AYDO;
AC 08AYDO;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Vascular endothelial cadherin precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
CX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Romero S., Jaffredo T., Dunon D.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF522067; AAN33002.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR Interpro; IPR002126; Cadherin.
DR Interpro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN-.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS02068; CADHERIN_2; 5.
DR KEGG; Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 773 AA; 86842 MW; ACS9CEH1119043F4 CRC64;

Query Match
Best Local Similarity 83.3%; Pred. No. 0.085;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DWIMNOMHIDEKN 12
Db 41 DWIMNOMHIDEKN 52

RESULT 3
Q01321 PRELIMINARY; PRT; 524 AA.
ID 001321;
AC 001321; P91858;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DR F26H9.1.
DR F26H9.1.
GN F26H9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

```

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Baynes C.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Gaxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kireten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons N., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Straden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman F.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Sims M.;
RL EMBL; 281516; CAB04206.1; -.
DR EMBL; 275535; CAB04206.1; JOINED.
DR EMBL; 275535; CAB04206.1; JOINED.
DR EMBL; 281516; CAB9829.1; JOINED.
DR PIR; T20872; T20872.
DR WormPeP; F26H9.1; CE09706.
DR Interpro; IPR008945; SKP1_Skp2.
SQ SEQUENCE 524 AA; 60462 MW; 6A85B0C0D0D2B76 CRC64;

Query Match
Best Local Similarity 54.3%; Score 50; DB 5; Length 524;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMNOMHIDEKN 15
Db 486 WIMNOMHIDEKN 499

RESULT 4
Q8C9N7 PRELIMINARY; PRT; 247 AA.
ID 08C9N7;
AC 08C9N7;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Cadherin 8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK041720; BAC31041.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR Interpro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 2.
DR PRINTS; PR00205; CADHERIN-.
DR SMART; SM00112; CA; 1.

```

DR PROSITE; PS00232; CADHERIN\_1; 1.  
DR PROSITE; PS50268; CADHERIN\_2; 2.  
SQ SEQUENCE 247 AA; 27565 MW; DF7551A97A29289D CRC64;

Query Match 53.3%; Score 49; DB 11; Length 247;  
Best Local Similarity 63.6%; Pred. No. 6.7;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 WINNOMHIDE 12  
Db 63 WYNNOMFVLEE 73

RESULT 5  
O8C449 PRELIMINARY; PRT; 716 AA.  
AC O8C449;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Cadherin 8.  
GN CDH8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Hippocampus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team.  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
DR EMBL; AK083092; BAC38758.1; -.  
DR MGD; MGI:107434; Cdh8.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.  
DR InterPro; IPR000233; Cadherin\_C\_term.  
DR Pfam; PF01049; Cadherin\_C\_term; 1.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SM00112; CA; 5.  
DR PROSITE; PS00232; CADHERIN\_1; 3.  
DR PROSITE; PS50268; CADHERIN\_2; 5.  
SQ SEQUENCE 716 AA; 79145 MW; 976F5D845F938BD CRC64;

Query Match 53.3%; Score 49; DB 11; Length 716;  
Best Local Similarity 63.6%; Pred. No. 21;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 WINNOMHIDE 12  
Db 63 WYNNOMFVLEE 73

RESULT 6  
O8C375 PRELIMINARY; PRT; 716 AA.  
AC O8C375;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Cadherin 8.  
GN CDH8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]

Query Match 53.3%; Score 49; DB 11; Length 754;  
Best Local Similarity 63.6%; Pred. No. 22;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team.  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
DR EMBL; AK086711; BAC39724.1; -.  
DR MGD; MGI:107434; Cdh8.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.  
DR InterPro; IPR000233; Cadherin\_C\_term.  
DR Pfam; PF00028; cadherin; 5.  
DR PRINTS; PR01049; Cadherin\_C\_term; 1.  
DR SMART; SM00112; CA; 5.  
DR PROSITE; PS00232; CADHERIN\_1; 3.  
DR PROSITE; PS50268; CADHERIN\_2; 5.  
SQ SEQUENCE 716 AA; 79149 MW; 5BBD598F6638624A CRC64;

Query Match 53.3%; Score 49; DB 11; Length 716;  
Best Local Similarity 63.6%; Pred. No. 21;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 WINNOMHIDE 12  
Db 63 WYNNOMFVLEE 73

RESULT 7  
O8BRK4 PRELIMINARY; PRT; 754 AA.  
AC O8BRK4;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Cadherin 8.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Cortex;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team.  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
DR EMBL; AK044046; BAC31751.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.  
DR InterPro; IPR000233; Cadherin\_C\_term.  
DR Pfam; PF00028; cadherin; 5.  
DR PRINTS; PR01049; Cadherin\_C\_term; 1.  
DR SMART; SM00112; CA; 5.  
DR PROSITE; PS00232; CADHERIN\_1; 3.  
DR PROSITE; PS50268; CADHERIN\_2; 5.  
SQ SEQUENCE 754 AA; 83624 MW; 149C3C70D3F21568 CRC64;

Query Match 53.3%; Score 49; DB 11; Length 754;  
Best Local Similarity 63.6%; Pred. No. 22;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;



## RESULT 11

Q87GE0 PRELIMINARY; PRT; 1497 AA.  
 AC O87GE0;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN VPAL376.  
 OS Vibrio parahaemolyticus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KMD 221063 / Serotype O3:K6;  
 RX MEDLINE=22508454; PubMed=12620739;  
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagamori K.,  
 RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
 RA Yasunaga T., Honda T., Shingawa H., Hattori M., Iida T.;  
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
 distance from that of V. cholerae.";  
 RL Lancet 361:743-749 (2003).  
 DR EMBL; AP005088; BAC62719.1;  
 DR InterPro; IPR000437; Prok 11proprot S.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 1497 AA; 165870 MW; 7EFB28D11BD8DB CRC64;

Query Match 50.0%; Score 46; DB 16; Length 1497;  
 Best Local Similarity 43.8%; Pred. No. 1,4e+02;  
 Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 WTWNQMH--IDBKNE 15  
 Db 935 WLMNDRYSYDQEKND 950

## RESULT 12

Q9YV168 PRELIMINARY; PRT; 143 AA.  
 AC Q9YV168;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE CG14611 protein.  
 GN PERI OR BCDNA:RH30329 OR CG10066 OR CG14611.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie J.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman W.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Barker E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertone P., Bhandari D., Bolshakov S.,  
 RA Butts K.C., Busam D.A., Butler H., Brokstein P., Brotlier P.,  
 RA Butts K.C., Busam D.A., Butler H., Brokstein P., Brotlier P.,  
 RA Chertey J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evaragelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Morkulov G., Mlshina N.V., Mobarry C., Morris J., Mostreli A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wastaman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195 (2000).  
 DR EMBL; AE003672; AAF54057.1;  
 DR Flybase; FBgn0037475; Ferl.  
 SQ SEQUENCE 143 AA; 16081 MW; 6A8B602464AFAD7E CRC64;

Query Match 49.5%; Score 45.5; DB 5; Length 143;  
 Best Local Similarity 53.3%; Pred. No. 14;  
 Matches 8; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

QY 1 DWIWNQMHIDBKNE 15  
 Db 39 DWIWSQ---EQEKEE 50

## RESULT 13

Q8RYF1 PRELIMINARY; PRT; 176 AA.  
 AC Q8RYF1;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative resistance protein (Fragment).  
 GN Avena.  
 OS Avena striigosa (black oat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Avenae; Avena.  
 OX NCBI\_TaxID=38783;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Loarce Y., Iritoyen M.L., Fominaya A., Ferrer E.,  
 RT "Resistance gene analogs in Avena striigosa".  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ437574; CAD26862.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0006915; P:Apoptosis; IEA.  
 DR InterPro; IPR002182; NB-ARC.  
 DR Pfam; PF00931; NB-ARC; 1.  
 FT NON TER 1  
 FT NON TER 1  
 SQ SEQUENCE 176 AA; 20133 MW; D60982A15C6FE2FE CRC64;

Query Match 48.9%; Score 45; DB 10; Length 176;  
 Best Local Similarity 58.3%; Pred. No. 21;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 TWNQMHIDBKNE 14  
 Db 62 LWKSMNIDBKND 73

## RESULT 14

Q86T88

```

ID Q86TS8 PRELIMINARY; PRT; 241 AA.
AC Q86TS8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Human full-length cDNA 5'-PRIME end of clone CSDB010YVP19 of placenta
DE of Homo sapiens (Human) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Genoscope;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Li W.B., Gruber C., Jessee J., Polayes D.;
RT "Full-length cDNA libraries and normalization";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX248750; CAD66557.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 2.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SMO0112; CA; 2.
DR PROSITE; PS00232; CADHERIN_1; 1.
DR PROSITE; PS50268; CADHERIN_2; 2.
KW Plasmid.
FT NON TER 241
SQ SEQUENCE 241 AA; 26348 MW; 4649831B55424604 CRC64;

Query Match 48.9%; Score 45; DB 4; Length 241;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WINNQHIDE 12
Db 46 WINNQFVIER 56

RESULT 15
Q8IK72 PRELIMINARY; PRT; 241 AA.
AC Q8IK72;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Hypothetical protein.
DE PF14_0734.
GN Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Perera M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA Mefraden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum.";
```

```

RL Nature 419:498-511 (2002).
DR EMBL; AE014828; AAN37347.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR008266; Tyr_Kinase_AS.
DR Pfam; PF00069; Kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Hypothetical protein.
FT NON TER 241
SQ SEQUENCE 241 AA; 28945 MW; 03664AF8F322F246 CRC64;

Query Match 48.9%; Score 45; DB 5; Length 241;
Best Local Similarity 44.4%; Pred. No. 29;
Matches 8; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 2 WINNQHIDE---DEKNE 15
Db 146 WINNQHILMDCSDPSKDE 163
```

Search completed: July 29, 2004, 11:37:09  
Job time : 32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:31:25 ; Search time 10.3333 Seconds

(Without alignments)  
139.633 Million cell updates/sec

Title: US-10-040-128-2

Sequence: 1 DWIMQMHIIDEKNE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	87	94.6	784	1	IJHUC5
2	50	53.3	524	2	T20872
3	49	53.3	793	2	D38992
4	47.5	51.6	388	1	D70006
5	45	48.9	391	2	C69673
6	45	48.9	796	2	A38992
7	45	48.9	796	2	I48277
8	45	48.9	796	2	I48556
9	45	48.9	796	2	A53584
10	44.5	48.4	963	2	T09911
11	44	47.8	163	2	F89999
12	44	47.8	558	2	E89962
13	44	47.8	558	2	T29503
14	44	47.8	785	2	I50180
15	44	47.8	790	2	I51638
16	44	47.8	790	2	I51638
17	44	47.8	790	2	I51638
18	43	46.7	435	2	C84941
19	42	45.7	77	2	A11306
20	42	45.7	82	2	T46510
21	42	45.7	296	2	F95380
22	42	45.7	315	2	AG3535
23	42	45.7	427	2	AF0590
24	41	44.6	77	2	A11678
25	41	44.6	314	2	S47696
26	41	44.6	314	2	H91171
27	41	44.6	314	2	H86017
28	41	44.6	431	2	B82119
29	41	44.6	519	2	T33616

30	41	44.6	532	2	G90607	hypothetical prote
31	41	44.6	712	2	T16338	hypothetical prote
32	41	44.6	789	2	I52701	K-cadherin - rat
33	41	44.6	790	2	I37016	cadherin-6 - human
34	41	44.6	790	2	I50178	cadherin-6B - chic
35	41	44.6	2241	2	T20971	hypothetical prote
36	41	44.6	2261	2	T20978	hypothetical prote
37	41	44.6	6831	2	A88852	protein unc-22 (lim
38	41	44.6	6839	2	S57242	twlctnln [similari
39	41	44.6	7160	2	T27935	hypothetical prote
40	40.5	44.0	479	2	T23643	hypothetical prote
41	40	43.5	120	2	C72425	hypothetical prote
42	40	43.5	308	2	AB1957	hypothetical prote
43	40	43.5	308	2	G75264	probable lipase -
44	40	43.5	355	2	T06122	cysteine proteinas
45	40	43.5	373	2	S48496	saccharopine dehyd

## ALIGNMENTS

### RESULT 1

IJHUC5

Cadherin 5 precursor - human  
N/Alternate names: 7B4 antigen; cadherin, endothelial-specific; VE-cadherin  
C/Species: Homo sapiens (man)

C/Date: 30-Jun-1993 #sequence revision 13-Sep-1996 #ext\_change 22-Jun-1999  
C/Accession: S49893; S24305; A43418

R/Breviario, F.; Cavada, L.; Corada, M.; Martin-Padura, I.; Golay, J.; Introna, M.; Lamp  
submitted to the EMBL Data Library, June 1994

A/Description: Molecular and functional properties of VE-cadherin (7B4/cadherin-5) a nov  
A/Reference number: S49893

A/Accession: S49893

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-784 <BRE>

A/Cross-references: EMBL:X79981; NID:G599833; PIDN:CA56306.1; PID:G599834  
R/Suzuki, S.; Sano, K.; Tanihara, H.

Cell Regul. 2, 261-270, 1991

A/Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous t  
A/Reference number: S24305; MUID:91283540; PMID:2059658

A/Accession: S24305

A/Molecule type: mRNA

A/Residues: 5-516, 'I', 518-784 <SUZ>

A/Cross-references: EMBL:X59796; NID:G639976; PIDN:CA42468.1; PID:G29593  
R/Lampugnani, M.G.; Resnati, M.; Raiteri, M.; Pigott, R.; Pisacane, A.; Hohen, G.; Ruco,

J. Cell Biol. 118, 1511-1522, 1992

A/Title: A novel endothelial-specific membrane protein is a marker of cell-cell contacts  
A/Reference number: A43418; MUID:92394977; PMID:1522121

A/Accession: A43418

A/Molecule type: protein

A/Residues: 48-60, 'X', 62, 'X', 64, 108-116, 'X', 118-123, 237-238, 'X', 240, 'X', 242-252, 'X', 254-  
A/Experimental source: cultured endothelial cells

A/Note: sequence extracted from NCBI backbone (NCBI:113040, NCBI:113045, NCBI:113047,  
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought t

C/Genetics:

A/Gene: GDB:CDH5

A/Cross-references: GDB:134230; OMIM:601120

A/Map position: 16q22.1-16q22.1

C/Superfamily: cadherin; cadherin repeat homology  
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro

F/1-25/Domain: signal sequence #status predicted <SIG>

F/26-47/Domain: propeptide #status predicted <PRO>

F/48-784/Product: cadherin 5 #status predicted <MAT>

F/48-593/Domain: extracellular #status predicted <EXT>

F/50-151/Domain: cadherin repeat homology <CR1>

F/154-258/Domain: cadherin repeat homology <CR2>

F/261-372/Domain: cadherin repeat homology <CR3>

F/375-479/Domain: cadherin repeat homology <CR4>

F/481-587/Domain: cadherin repeat homology <CR5>

F/594-620/Domain: transmembrane #status predicted <TM>

F/621-784/Domain: intracellular #status predicted <INT>

F:61,112,157,362,442,523,535/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 94.6%; Score 87; DB 1; Length 784;  
Best Local Similarity 100.0%; Pred. No. 7,6e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNE 14  
|||||  
DB 48 DWIMQMHIDEKNE 61

## RESULT 2

hypothetical protein F26H9.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T20872; T21443

R: Sims, M.  
Submitted to the EMBL Data Library, June 1996

A:Reference number: Z19337

A:Accession: T20872

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-524 <W1L>

A:Cross-references: EMBL:Z75535; PIDN:CAA9829.1; GSPDB:GN00019; CESP:F26H9.1

A:Experimental source: clone F14B4

R: Baynes, C.  
Submitted to the EMBL Data Library, November 1996

A:Reference number: Z19422

A:Accession: T21443

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-524 <W12>

A:Cross-references: EMBL:Z81516; PIDN:CAB04206.1; GSPDB:GN00019; CESP:F26H9.1

A:Experimental source: clone F26H9

C:Genetics:

A:Gene: CESP:F26H9.1

A:Map position: 1

A:introns: 49/2; 99/1; 136/1; 186/1; 228/3; 287/3; 346/1; 403/1; 449/3

Query Match 54.3%; Score 50; DB 2; Length 524;  
Best Local Similarity 50.0%; Pred. No. 3.8;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMQMHIDEKNE 15  
|||  
DB 486 WDMQMHIDEKNE 499

## RESULT 3

cadherin 8 - human  
C:Species: *Homo sapiens* (man)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jan-2000  
C:Accession: D38992

R: Suzuki, S.; Sano, K.; Tanihara, H.  
Cell Regul. 2, 261-270, 1991

A:Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous

A:Reference number: S24305; WUID:91283540; PMID:2059658

A:Accession: D38992

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-793 <SUZ>

A:Cross-references: GB:L34060; NID:G506411; PIDN:AAA35628.1; PID:G506412

C:Genetics:

A:Gene: GDB:CDH8

A:Cross-references: GDB:5822911

C:Superfamily: cadherin repeat homology

C:Keywords: calcium binding; cell adhesion; duplication

F:163-269/Domain: cadherin repeat homology <CDH>

Query Match 53.3%; Score 49; DB 2; Length 793;  
Best Local Similarity 63.6%; Pred. No. 8.7;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WIMQMHIDE 12  
|||  
DB 56 WIMQMHIDE 66

## RESULT 4

conserved hypothetical protein yuba - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C:Accession: D70006

R: Kunst, F.; Ogatawa, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritzt, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogatawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scottone, P.; Sekiguchi, J.; Sekowska, A.; Seror,

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpete, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumslein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; WUID:98044033; PMID:9384377

A:Accession: D70006

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-388 <KUN>

A:Cross-references: GB:Z99119; GB:AL009126; NID:G2635411; PIDN:CAB15094.1; PID:G2635600

A:Experimental source: strain 168

C:Genetics:

A:Gene: yuba

C:Superfamily: *Bacillus subtilis* conserved hypothetical protein yuef

Query Match 51.6%; Score 47.5; DB 1; Length 388;  
Best Local Similarity 62.5%; Pred. No. 6.8;  
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 DWIMQMHIDEKNE 15  
|||  
DB 366 DWIMQMHIDEKNE 381

## RESULT 5

penicillin-binding protein pbpx - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999

C:Accession: C69673

R: Kunst, F.; Ogatawa, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritzt, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogatawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scottone, P.; Sekiguchi, J.; Sekowska, A.; Seror,

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpete, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumslein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; WUID:98044033; PMID:9384377

A:Accession: C69673

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA



A;Residues: 1-391 <KUN>  
A;Cross-references: GB:Z29112; GB:AL009126; NID:g2633902; PIDN:CAB13568.1; PID:e1185286;  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: pbpx

Query Match	48.9%	Score	45	DB	2	Length	391
Best Local	Similarity	Pred	53.8%		17		
Matches	7	Conservative	3	Mismatches	3	Indels	0
						Gaps	0

RESULT 6  
A38992  
catherin 11 precursor - human  
N:Alternate names: OB-catherin, osteoblast  
C/Species: Homo sapiens (man)  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 20-Aug-1999  
C/Accession: A38992  
R/Suzuki, S.; Sano, K.; Tanihara, H.  
Cell Regul. 2, 261-270, 1991  
A/Title: Diversity of the catherin family: evidence for eight new catherins in nervous t  
A/Reference numbers: S24305; MUID:91283540; PMID:2059658  
Accession: 138992

Query Match	48.9%	Score 45;	DB 2;	Length 796;
Best Local Similarity	54.5%	Pred. No. 38;		
Matches	6;	Conservative	2;	Mismatches 3;
				Indels 0;
				Gaps 0;

RESULT 7  
148277  
cadherin-11 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jan-2000  
C:Accession: I48277  
R:Hoffmann, I.; Balling, R.  
Dev. Biol. 169, 337-346, 1995  
A>Title: Cloning and expression analysis of a novel mesodermally expressed cadherin  
A:Reference number: I48277; MUID:95269886; PMID:7750649  
A:Accession: I48277  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-796 <RES>  
A:Cross-references: EMBL:X77557; NID:g642796; PIDD:CAAB4674.1; PID:g666071  
C:Genetics:  
A:Gene: cad-11  
C:Superfamily: cadherin; cadherin repeat homology  
C:56-159/Domain: cadherin repeat homology <CDH>

Query Match	48.9%;	Score 45;	DB 2;	Length 796;
Best Local Similarity	54.5%;	Pred. No. 38;		

	Matches	6;	Conservative	2;	Mismatches	3;	Indels	0;	Gaps	0;
Qy	2	WINQMHIDE	12	: : :	:					
Db	55	WVWNGFFVIEE	65							

```

RESULT 8
I49556
C:cadherin-11 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: I49556
R:Kimura, Y.; Matsumami, H.; Inoue, T.; Shimamura, K.; Uchida, N.; Ueno, T.; Miyazaki, T.
Dev. Biol. 169, 347-358, 1995
A:Title: Cadherin-11 expressed in association with mesenchymal morphogenesis in the heart
A:Reference number: I49556; M01D:95269887; PMID:7750650
A:Accession: I49556
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residue: 1-796 <RES>
A:Cross-references: GB:031963; NID:g974190; P1DN:BAA06730.1; PID:g974191
C:Superfamily: cadherin, cadherin repeat homology
F:55-159/Domain: cadherin repeat homology <CDH>

```

Query Match	48.9%	Score 45;	DB 2;	Length 796;
Best Local Similarity	54.5%	Pred. NO. 38;		
Matches 6;	Conservative 2;	Mismatches 3;	Indels 0;	Gaps 0;

```

RESULT 9
A:53584
OB-cadherin precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 07-Oct-1994 #sequence_rev1sion 07-Oct-1994 #ext_change 20-Jun-2000
C:Accession: A53584
R:Okazaki, M.; Takehisa, S.; Kawai, S.; Kikuno, R.; Tsujimura, A.; Kudo, A.; Mann, E.
J. Biol. Chem. 269, 12092-12098, 1994
A:Title: Molecular cloning and characterization of OB-cadherin, a new member of cadherin
A:Reference number: A53584; MUID:94216322; PMID:816313
A:Accession: A53584
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-796 <OKA>
A:Cross-references: GB:ID12553; NID:9994774; PIDN:BAA04797.1; PID:9994775
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: transmembrane protein
F:56-159/Domain: cadherin repeat homology <CR1>
F:162-268/Domain: cadherin repeat homology <CR2>
F:386-488/Domain: cadherin repeat homology <CR4>

```

Query Match	48.9%	Score 45;	DB 2;	Length 796;
Best Local Similarity	54.5%;	Pred. No. 38;		
Matches	6; Conservative	2; Mismatches	3; Indels	0; Gaps

RESULT 10  
 T09911  
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) T22A6\_310 - Arabidopsis thaliana (mouse-ear cress)  
 C.Species: Arabidopsis thaliana (mouse-ear cress)  
 C.Date: 16-Jul-1999 #sequence revision 16-Jul-1999 #text\_change 15-Oct-1999  
 C.Accession: T09911  
 R.Bevan, M.; Zimmermann, W.; Grunewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, A.; Reference number: Z16896  
 submitted to the Protein Sequence Database, June 1999



R;Espeseth, A.; Johnson, E.; Kintner, C.  
 Mol. Cell. Neurosci. 6, 199-212, 1995  
 A;Title: Xenopus F-cadherin, a novel member of the cadherin family of cell adhesion mole  
 A;Reference number: 151638; MUID:96039533; PMID:7496627  
 A;Accession: 151638  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-790 <ESP>  
 A;Cross-references: EMBL:X85330; NID:9854634; PIDD:CAA59679.1; PID:9854635  
 C;Superfamily: cadherin; cadherin repeat homology  
 F;161-267/Domain: cadherin repeat homology <CR2>

Query Match 47.8%; Score 44; DB 2; Length 790;  
 Best Local Similarity 54.5%; Pred. No. 54;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WINQMHIDE 12  
 | : | | | : | |  
 Db 54 WVNQFFVLEE 64

Search completed: July 29, 2004, 11:37:50  
 Job time : 11.333 secs

***This Page Blank (uspto)***



CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas, The  
CC adenocarcinomas, adenomas, leukaemic tumours or lymphoid tumours). The  
CC composition or antibody is also useful for treating a cell proliferation  
CC disorder associated with vascularisation (e.g. blood vessel proliferation  
CC disorder, fibrotic disorders, angiogenesis, tumour growth, tumour  
CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
CC These may also be used for reducing or inhibiting tumour vasculature in a  
CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
CC therapy, particularly for inhibiting angiogenesis or tumour  
CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
CC domain 1 antibody binding peptide  
CC  
XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 92; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 2e-07; Mismatches 0; Gaps 0;

Matches 15; Conservative 0; Indels 0; Gaps 0;

QY 1 DWIMNQHIIDEKKE 15  
Db 1 DWIMNQHIIDEKKE 15

RESULT 2  
AAU08786 standard; protein; 103 AA.

AAU08786;  
27-FEB-2002 (first entry)

DE Murine VE-cadherin N-terminal domain 1 polypeptide.

XX VE-cadherin; angiogenesis; neoplastic disease;  
XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
XX paracellular permeability; solid tumour; collagenous vascular disease;  
XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
XX carcinoma; sarcoma; adenocarcinoma; adenomas; gene therapy;  
XX leukaemic tumour; lymphoid tumours; cell proliferation disorder;  
XX vascularisation; blood vessel proliferation disorder; fibrotic disorder;  
XX age-related muscular degeneration; tumour neovascularisation; mouse;  
XX VE-cadherin N-terminal domain 1.

OS Mus sp.

PN WO200175109-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010505.

PR 31-MAR-2000; 2000US-00540967.

PA (IMCL-) IMCLONE SYSTEMS INC.

PI Liao F, Hicklin DJ, Bohlen P;

DR WPI; 2001-656988/75.

XX New antibody antagonists of VE-cadherin, which does not adversely affect  
XX PT vascular permeability, useful for inhibiting angiogenesis or tumour  
XX PT metastasis, e.g. autoimmune disease, carcinomas or leukaemic tumors.

XX Example 2; Fig 2; 44p; English.

XX The method relates to an antibody or an antibody fragment, capable of

CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
CC mediated adherens junction formation in vitro, but does not exert any  
CC significant or substantial effect on paracellular permeability in vitro.  
CC The antibody or antibody fragment is capable of specifically binding to a  
CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas, The  
CC adenocarcinomas, adenomas, leukaemic tumours or lymphoid tumours). The  
CC composition or antibody is also useful for treating a cell proliferation  
CC disorder associated with vascularisation (e.g. blood vessel proliferation  
CC disorder, fibrotic disorders, angiogenesis, tumour growth, tumour  
CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
CC These may also be used for reducing or inhibiting tumour vasculature in a  
CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
CC therapy, particularly for inhibiting angiogenesis or tumour  
CC neovascularisation. This sequence represents a murine VE-cadherin N-  
CC terminal domain 1 polypeptide  
CC  
XX

SQ Sequence 103 AA;

Query Match 100.0%; Score 92; DB 4; Length 103;

Best Local Similarity 100.0%; Pred. No. 1.7e-06; Mismatches 0; Gaps 0;

Matches 15; Conservative 0; Indels 0; Gaps 0;

QY 1 DWIMNQHIIDEKKE 15  
Db 1 DWIMNQHIIDEKKE 15

RESULT 3  
AAU08779 standard; peptide; 15 AA.

AAU08779;  
27-FEB-2002 (first entry)

DE VE-cadherin N-terminal domain 1 antibody binding peptide #3.

XX VE-cadherin; antibody binding peptide; angiogenesis; neoplastic disease;  
XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
XX paracellular permeability; solid tumour; collagenous vascular disease;  
XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
XX carcinoma; sarcoma; adenocarcinoma; adenomas; gene therapy;  
XX leukaemic tumour; lymphoid tumours; cell proliferation disorder;  
XX vascularisation; blood vessel proliferation disorder; fibrotic disorder;  
XX age-related muscular degeneration; tumour neovascularisation;  
XX VE-cadherin N-terminal domain 1.

OS Mammalia.

PN WO200175109-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010505.

PR 31-MAR-2000; 2000US-00540967.

PA (IMCL-) IMCLONE SYSTEMS INC.

PI Liao F, Hicklin DJ, Bohlen P;

DR WPI; 2001-656988/75.

XX New antibody antagonists of VE-cadherin, which does not adversely affect  
XX PT vascular permeability, useful for inhibiting angiogenesis or tumour  
XX PT metastasis, e.g. autoimmune disease, carcinomas or leukaemic tumors.

XX Claim 1; Page 32; 44pp; English.  
 XX  
 PS The method relates to an antibody or an antibody fragment, capable of  
 CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 CC mediated adherens junction formation in vitro, but does not exert any  
 CC significant or substantial effect on paracellular permeability in vitro.  
 CC The antibody or antibody fragment is capable of specifically binding to a  
 CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis in an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 CC adenosarcomas, adenomas, leukaemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vascularisation in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
 CC domain 1 antibody binding peptide  
 XX  
 SQ Sequence 15 AA;  
 Query Match 94.6%; Score 87; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DWINQMHIDEKN 14  
 Db 1 DWINQMHIDEKN 14  
 RESULT 4  
 AAU08785  
 ID AAU08785 standard; protein; 104 AA.  
 XX  
 AC AAU08785;  
 XX  
 DT 27-FEB-2002 (first entry)  
 XX  
 DE Human VE-cadherin N-terminal domain 1 polypeptide.  
 XX  
 KW VE-cadherin; angiogenesis; neoplastic disease;  
 KW VE-cadherin mediated adherens junction formation; autoimmune disease;  
 KW paracellular permeability; solid tumour; collagenous vascular disease;  
 KW rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
 KW retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
 KW leukaemic sarcoma; adenocarcinoma; adenosarcoma; adenoma; gene therapy;  
 KW vascularisation; lymphoid tumours; cell proliferation disorder;  
 KW age-related muscular degeneration; tumour neovascularisation; human;  
 KW VE-cadherin N-terminal domain 1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175109-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US010505.  
 XX  
 PR 31-MAR-2000; 2000US-00540967.  
 XX  
 PA (IMCL-) IMCLONE SYSTEMS INC.  
 XX  
 PI Liao F, Hicklin DJ, Bohlen P;  
 XX  
 DR WPI; 2001-656986/75.

XX New antibody antagonists of VE-cadherin, which does not adversely affect  
 PT vascular permeability, useful for inhibiting angiogenesis or tumor  
 PT metastasis, e.g. autoimmune disease, carcinomas or leukaemic tumors.  
 XX  
 PS Example 2; Fig 2; 44pp; English.  
 XX  
 CC The method relates to an antibody or an antibody fragment, capable of  
 CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 CC mediated adherens junction formation in vitro, but does not exert any  
 CC significant or substantial effect on paracellular permeability in vitro.  
 CC The antibody or antibody fragment is capable of specifically binding to a  
 CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis in an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 CC adenosarcomas, adenomas, leukaemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vascularisation in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a human VE-cadherin N-  
 CC terminal domain 1 polypeptide  
 XX  
 SQ Sequence 104 AA;  
 Query Match 94.6%; Score 87; DB 4; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DWINQMHIDEKN 14  
 Db 1 DWINQMHIDEKN 14  
 RESULT 5  
 AAY64578  
 ID AAY64578 standard; peptide; 108 AA.  
 XX  
 AC AAY64578;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Nonclassical cadherin extracellular domain SEQ ID NO:6.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; Li-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease.  
 XX  
 OS Mammalia.  
 XX  
 PN WO9957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 05-MAY-1999; 99WO-CA000363.  
 XX  
 PR 05-MAY-1998; 98US-00073040.  
 XX  
 PR 06-NOV-1998; 98US-00187859.  
 XX  
 PR 20-JAN-1999; 99US-00234395.  
 XX  
 PR 08-MAR-1999; 99US-00264516.

PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX Blaschuk OW, Gour BJ, Byers SJ  
 PI  
 XX  
 XX  
 DR MPI; 2000-038791/03.  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT mediated functions for treating e.g. cancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX  
 PS Disclosure; Fig 2; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MA can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
 CC -related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in bioreactors.  
 CC AAY60592 to AAY64572 represent specifically claimed peptides, and  
 CC AAY64573 to AAY64643 and AA23183 to AA233186 represent sequences used in  
 CC the exemplification of the present invention  
 XX  
 SQ Sequence 108 AA;  
 Query Match 94.6%; Score 87; DB 3; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 1e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DWIMNQMHIIDEKN 14  
 Db 5 DWIMNQMHIIDEKN 18  
 RESULT 6  
 AAU11276  
 ID AAU11276 standard; protein; 566 AA.  
 XX  
 AC AAU11276;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human recombinant KVE702 polypeptide.  
 XX  
 KM Human; KV protein; KVE702; angiogenic factor antagonist; angiogenesis;  
 KM vascular endothelial structure regulator; cancer; vascular endothelium;  
 KM tumour; blood vessel; epithelial cell; cytostatic; gene therapy.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 PN WO200183562-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 04-MAY-2001; 2001WO-GB001956.  
 XX  
 PR 04-MAY-2000; 2000GB-00010630.  
 XX  
 PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.

XX  
 XX Jiang WG;  
 PI  
 XX  
 XX MPI; 2002-066521/09.  
 DR  
 DR N-PSDB; AAS16965.  
 XX  
 PT Novel nucleic acid for inhibition of angiogenesis and prophylaxis or  
 PT treatment of cancer, comprises both an angiogenic factor antagonist and a  
 PT vascular endothelial structure regulator.  
 XX  
 PS Disclosure; Fig 2; 66pp; English.  
 XX  
 CC The invention relates to recombinant nucleic acid sequences that encode  
 CC both an angiogenic factor antagonist and a vascular endothelial structure  
 CC regulator, collectively referred to as a KV protein. KV DNA and proteins,  
 CC and also cells and compositions containing the sequences, are useful in  
 CC treating or preventing cancer or angiogenesis in mammals, in particular  
 CC humans. KV proteins are also useful for regulation of the development of  
 CC blood vessels and their formation, in the vascular endothelium and/or  
 CC tumour. The KVE702 gene and its fragments are useful in transfection of  
 CC human epithelial cells and to generate products suitable for angiogenesis  
 CC intervention. This sequence represents the human KVE702 protein  
 XX  
 SQ Sequence 566 AA;  
 Query Match 94.6%; Score 87; DB 5; Length 566;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DWIMNQMHIIDEKN 14  
 Db 474 DWIMNQMHIIDEKN 487  
 RESULT 7  
 AAU11279  
 ID AAU11279 standard; protein; 594 AA.  
 XX  
 AC AAU11279;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human recombinant U37 polypeptide.  
 XX  
 KM Human; KV protein; KVE702; angiogenic factor antagonist; angiogenesis;  
 KM vascular endothelial structure regulator; cancer; vascular endothelium;  
 KM tumour; blood vessel; epithelial cell; cytostatic; gene therapy; U37.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 PN WO200183562-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 04-MAY-2001; 2001WO-GB001956.  
 XX  
 PR 04-MAY-2000; 2000GB-00010630.  
 XX  
 PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.  
 XX  
 PI Jiang WG;  
 XX  
 DR MPI; 2002-066521/09.  
 DR  
 DR N-PSDB; AAS16975.  
 XX  
 PT Novel nucleic acid for inhibition of angiogenesis and prophylaxis or  
 PT treatment of cancer, comprises both an angiogenic factor antagonist and a  
 PT vascular endothelial structure regulator.



XX PS Disclosure; Fig 15; 66pp; English.  
 XX CC The invention relates to recombinant nucleic acid sequences that encode  
 CC both an angiogenic factor antagonist and a vascular endothelial structure  
 CC regulator, collectively referred to as a KV protein. KV DNA and proteins,  
 CC and also cells and compositions containing the sequences, are useful in  
 CC treating or preventing cancer or angiogenesis in mammals, in particular  
 CC humans. KV proteins are also useful for regulation of the development of  
 CC blood vessels and their formation, in the vascular endothelium and/or  
 CC tumour. The KVE702 gene and its fragments are useful in transfection of  
 CC human epithelial cells and to generate products suitable for angiogenesis  
 CC intervention. This sequence represents the human J37 protein used in  
 CC methods of the invention  
 XX SQ Sequence 594 AA:  
 XX  
 QY Query Match 94.6%; Score 87; DB 5; Length 594;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 DWIMNQHIDEKN 14  
 49 DWIMNQHIDEKN 62  
 XX  
 RESULT 8  
 AAW25634  
 ID AAW25634 standard; protein; 780 AA.  
 XX AC AAW25634;  
 XX DT 25-MAR-2003 (revised)  
 DT 03-NOV-1997 (first entry)  
 XX DE Human cadherin-5.  
 XX KW Human; cadherin; rat; calcium-dependent cell adhesion protein;  
 KW superfamily; cytoskeleton; eatenin; cancer.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT Misc-difference 622  
 FT /note= "encoded by AAG"  
 XX PN US5646250-A.  
 PD 08-JUL-1997.  
 XX PF 01-NOV-1994; 94US-00332638.  
 XX PR 17-APR-1992; 92US-00872643.  
 PR 19-APR-1993; 93US-00049460.  
 XX PA (DOHE-) DOHENY EYE INST.  
 XX PI Suzuki S;  
 XX DR WPI; 1997-362997/33.  
 DR N-PSDB; AAT85401.  
 XX PT Human and rat cadherin polypeptide(s) - mediate cell-cell adhesion.  
 PS Claim 1; Col 69-72; 56pp; English.  
 XX CC This sequence represents human cadherin-5. The invention specifically  
 CC provides details of human cadherin-5, -8, -11, -12 and -13, and rat  
 CC cadherin-8, -5, -11 and -13. Cadherins are calcium-dependent cell  
 CC adhesion proteins. They are glycosylated integral membrane proteins that  
 CC have an N-terminal extracellular domain that determines binding  
 CC specificity, a hydrophobic membrane spanning region and a C-terminal  
 CC cytoplasmic domain, which is highly conserved among members of the

CC superfamily. The C-terminal domain interacts with the cytoskeleton  
 CC through eatenins and other cytoskeleton-associated proteins. The novel  
 CC cadherin proteins may be used in the analysis of the role of cadherins in  
 CC various cancers. Sequence analysis of the cadherin proteins also allows  
 CC investigation of the structure and function of cadherin. The cadherin  
 CC proteins may be isolated by using anti-cadherin antibodies. These  
 CC antibodies may also be used to modulate the activity of cadherin and to  
 CC determine the tissue specific distribution of cadherin proteins. Each  
 CC subclass of cadherins has a unique tissue distribution pattern. (updated  
 CC on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 780 AA:  
 XX  
 QY Query Match 94.6%; Score 87; DB 2; Length 780;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 DWIMNQHIDEKN 14  
 44 DWIMNQHIDEKN 57  
 XX  
 RESULT 9  
 AAW13130  
 ID AAW13130 standard; protein; 780 AA.  
 XX AC AAW13130;  
 XX DT 25-MAR-2003 (revised)  
 DT 13-MAY-1997 (first entry)  
 XX DE Full length human cadherin-5.  
 XX KW Ca2+ dependent; cell adhesion protein; placental; cadherin; rat; brain;  
 KW human; antibody; purification; determination; tissue expression;  
 KW binding antagonist; calcium ion.  
 XX OS Homo sapiens.  
 XX PN US5597725-A.  
 PD 28-JAN-1997.  
 XX PF 26-JAN-1994; 94US-00188228.  
 XX PR 17-APR-1992; 92US-00872643.  
 PR 19-APR-1993; 93US-00049460.  
 XX PA (DOHE-) DOHENY EYE INST.  
 XX PI Suzuki S;  
 XX DR WPI; 1997-108328/10.  
 DR N-PSDB; AAT61921.  
 XX PT Antibodies to cadherin proteins - useful as cadherin antagonists, etc.  
 PS Claim 5; Col 75-78; 59pp; English.  
 XX CC The present sequence is full length human cadherin-5, which is a Ca2+  
 CC dependent cell adhesion protein. The human cadherin cDNA was isolated  
 CC from a placental cDNA library, using probes based on homologous rat  
 CC cadherin cDNA. Antibodies or fragments that specifically bind the human  
 CC cadherin can be used to purify the cadherin, determine its tissue  
 CC expression and antagonise its ligand/antiligand binding activities.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 780 AA:  
 XX  
 QY Query Match 94.6%; Score 87; DB 2; Length 780;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWTNQMHIIDEKN 14  
 |||||  
 Db 44 DWTNQMHIIDEKN 57

## RESULT 10

ID ABB81477 standard; protein; 784 AA.

AC ABB81477;

DT 30-AUG-2002 (first entry)

DE Human VE-cadherin protein SEQ ID NO:52.

XX Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;

KM o-catenin; colon cancer.

OS Homo sapiens.

PN US2002045591-A1.

PD 18-APR-2002.

PF 17-JUL-2001; 2001US-00905983.

PR 26-MAY-1998; 98IL-00124650.

PR 26-MAY-1999; 99US-00318633.

PA (GEIG/) GEIGER B.

PA (BENZ/) BEN-ZE'EV A.

PA (SADO/) SADO E.

PI Geiger B, Ben-Ze'ev A, Sado E;

XX WPI; 2002-499105/53.

DR N-PSDB; AEN89394.

PT New construct encoding soluble cytoplasmic portion of cadherin including

PT beta catenin binding domain useful in treating cancer associated with

PT high beta-catenin activity e.g. colon cancer and melanoma.

XX Example 3; Page 56-58; 102pp; English.

CC The present invention describes a pharmaceutical composition for treating  
 CC cancer associated with abnormally high beta-catenin activity. The  
 CC pharmaceutical composition comprises a gene therapy vehicle harbouring a  
 CC polynucleotide that contains: (i) a nucleotide sequence encoding a  
 CC soluble cytoplasmic portion of a cadherin which lacks a transmembrane  
 CC portion and an extracellular portion of the cadherin, and includes a beta  
 CC -catenin binding domain; and (b) an upstream promoter for directing  
 CC expression of the soluble cytoplasmic portion of the cadherin in a  
 CC mammalian cell. Also described is a pharmaceutical composition for  
 CC treating cancer associated with abnormally high activity levels of beta-  
 CC catenin comprising a gene therapy vehicle harbouring a polynucleotide  
 CC that contains: (a) a nucleotide sequence encoding an o-catenin; and (b)  
 CC an upstream promoter for directing expression of the o-catenin in a  
 CC mammalian cell. The pharmaceutical compositions have cytostatic activity  
 CC and can be used in the suppression of beta-catenin-mediated  
 CC transactivation. They can be used for treating cancers associated with  
 CC abnormally high activity levels of beta-catenin such as colon cancers and  
 CC melanomas, by reducing these high activity levels of beta-catenin in  
 CC mammalian cells. The present sequence represents human VE-cadherin which  
 CC is used in the exemplification of the present invention

XX Sequence 784 AA;

Query Match 94.6%; Score 87; DB 5; Length 784;

Best Local Similarity 100.0%; Pred. No. 9.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWTNQMHIIDEKN 14  
 |||||

Db 48 DWTNQMHIIDEKN 61

## RESULT 11

ID ABU03507 standard; protein; 784 AA.

AC ABU03507;

DT 21-JAN-2003 (first entry)

DE Angiogenesis-associated human protein sequence #52.

XX Human; angiogenesis-associated transcript; angiogenesis;

KM angiogenesis-associated disease; cancer; cytostatic.

OS Homo sapiens.

PN WO200279492-A2.

PD 10-OCT-2002.

PF 14-FEB-2002; 2002MO-US004915.

PR 14-FEB-2001; 2001US-00784356.

PR 22-FEB-2001; 2001US-00791390.

PR 19-APR-2001; 2001US-0285475P.

PR 03-AUG-2001; 2001US-0310025P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334244P.

PA (BOSB-) EOS BIOTECHNOLOGY INC.

PA Murray R, Glynn R, Watson SR, Aziz N;

XX WPI; 2003-040681/03.

DR N-PSDB; ABX08791.

PT Detecting angiogenesis-associated transcript in a cell for diagnosing and

PT treating cancer by contacting a sample with a polynucleotide that

PT exhibits changes in expression level as a function of time in tissue

PT undergoing angiogenesis.

XX Example 2; Page 230; 291pp; English.

CC The present invention relates to methods and compositions for detecting  
 CC an angiogenesis-associated transcript in a cell in a patient. The method  
 CC involves contacting a biological sample from the patient with a  
 CC polynucleotide that selectively hybridizes to a sequence at least 80%  
 CC identical to any of the angiogenesis-associated human polynucleotide  
 CC sequences given in the specification. These angiogenesis-associated  
 CC polynucleotide sequences comprise genes that exhibit changes in  
 CC expression levels as a function of time in tissue undergoing  
 CC angiogenesis. The method and the polynucleotide sequences of the  
 CC invention are useful for diagnosing and treating angiogenesis and  
 CC angiogenesis-associated diseases e.g. cancer. The polynucleotide  
 CC sequences are also useful in the gene therapy of such disorders. The  
 CC angiogenesis-associated proteins encoded by the polynucleotide sequences  
 CC are useful as a vaccine for therapeutic and prophylactic immunisation.  
 CC ABU03456-ABU03569 represent angiogenesis-associated protein sequences

XX Sequence 784 AA;

Query Match 94.6%; Score 87; DB 6; Length 784;

Best Local Similarity 100.0%; Pred. No. 9.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWTNQMHIIDEKN 14  
 |||||

Db 48 DWTNQMHIIDEKN 61

RESULT 12



XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
 XX paracellular permeability; solid tumour; collagenous vascular disease;  
 XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
 XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
 XX carcinoma; sarcoma; adenocarcinoma; adenocarcinoma; adenoma; gene therapy;  
 XX leukemic tumour; lymphoid tumours; cell proliferation disorder;  
 XX vasculatization; blood vessel proliferation disorder; fibrotic disorder;  
 XX age-related muscular degeneration; tumour neovascularisation;  
 XX VE-cadherin N-terminal domain 1.  
 XX Mammalia.  
 XX WO200175109-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US010505.  
 XX 31-MAR-2000; 2000US-00540967.  
 XX (IMCL-) IMCLONE SYSTEMS INC.  
 XX Liao F, Hicklin DJ, Bohlen P;  
 XX WPI; 2001-656998/75.  
 XX New antibody antagonists of VE-cadherin, which does not adversely affect  
 XX PT vascular permeability, useful for inhibiting angiogenesis or tumor  
 XX PT metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.  
 XX Claim 1; Page 32; 44pp; English.  
 XX The method relates to an antibody or an antibody fragment, capable of  
 XX CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 XX CC mediated adherens junction formation in vitro, but does not exert any  
 XX CC significant or substantial effect on paracellular permeability in vitro.  
 XX CC The antibody or antibody fragment is capable of specifically binding to a  
 XX CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 XX CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 XX CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 XX CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 XX CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 XX CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 XX CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 XX CC adenocarcinomas, adenomas, leukemic tumours or lymphoid tumours). The  
 XX CC composition or antibody is also useful for treating a cell proliferation  
 XX CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 XX CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 XX CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 XX CC These may also be used for reducing or inhibiting tumour vasculature in a  
 XX CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 XX CC therapy, particularly for inhibiting angiogenesis or tumour  
 XX CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
 XX CC domain 1 antibody binding peptide  
 XX Sequence 15 AA:  
 XX  
 XX Query Match 84.8%; Score 78; DB 4; Length 15;  
 XX Best Local Similarity 93.3%; Pred. No. 2.8e-05;  
 XX Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX DE Novel human polypeptide #177.  
 XX Human; angiogenesis; cytokine; cell proliferation; pluripotent;  
 XX cell differentiation; totipotent; stem cell; transplantation; bio-sensor;  
 XX neuroepithelial cell; autoimmune disease; neural cell; genetic disorder;  
 XX nerve; brain tissue; central nervous system disease;  
 XX peripheral nerve system disease; neuropathy; haematopoiesis; bone;  
 XX myeloid disorder; lymphoid cell disorder; platelet disorder; tendon;  
 XX regeneration; cartilage; tendon; ligament; nerve tissue growth;  
 XX tissue repair; wound healing; burn; ulcer; osteoporosis; cancer;  
 XX osteoarthritis; bone degenerative disorder; periodontal disease;  
 XX gut protection; lung fibrosis; liver fibrosis; reperfusion injury;  
 XX immune deficiency; infection; autoimmune disorder; allergic reaction;  
 XX thrombolytic; thrombosis; coagulation disorder; hereditary disorder;  
 XX biorythm; circadian cycle; fertility; metabolism; catabolism; anabolism;  
 XX neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;  
 XX haemostatic; vulnery; antidiabetic; osteopathic; antiarthritic;  
 XX vasotropic; immunostimulant; antibacterial; fungicide; immunosuppressive;  
 XX antirheumatic; antidiabetic; antidiabetic; cytostatic; vitruide.  
 XX Homo sapiens.  
 XX WO2003023013-A2.  
 XX 20-MAR-2003.  
 XX 13-SEP-2002; 2002WO-US029001.  
 XX 13-SEP-2001; 2001US-032251P.  
 XX 12-SEP-2002; 2002US-00243552.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;  
 XX WPI; 2003-312249/30.  
 XX N-PSDB; ACD05667.  
 XX Novel nucleic acids and polypeptides for diagnosis, treatment of central  
 XX PT and peripheral nervous system diseases and neuropathies, such as  
 XX PT Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 XX PT lateral sclerosis.  
 XX Claim 20; SEQ ID NO 513; 300pp; English.  
 XX The present invention relates to the isolation of novel human  
 XX CC polynucleotide sequences and their encoding polypeptides. The novel  
 XX CC polypeptides exhibit activities relating to angiogenesis, cytokine, cell  
 XX CC proliferation, cell differentiation, antiinflammatory, and stem cell  
 XX CC growth factor activities. The polypeptides are involved in the  
 XX CC proliferation, differentiation and survival of pluripotent and totipotent  
 XX CC stem cells, and are useful for re-engineering damaged or diseased  
 XX CC tissues, transplantation, manufacture of bio-pharmaceuticals and  
 XX CC development of bio-sensors. The polypeptides can be used to manipulate  
 XX CC stem cells in culture to give rise to neuroepithelial cells that can be  
 XX CC used to augment or replace cells damaged by illness, autoimmune disease,  
 XX CC accidental damage or genetic disorders. The polypeptides induce the  
 XX CC proliferation of neural cells and regeneration of nerve and brain tissue  
 XX CC and are useful for the treatment of central and peripheral nervous system  
 XX CC diseases and neuropathies, such as Alzheimer's, Parkinson's disease,  
 XX CC Huntington's disease, amyotrophic lateral sclerosis (ALS). The  
 XX CC polypeptides are also involved in chemotactic or chemokinetic activity,  
 XX CC regulation of haematopoiesis and are useful for treating myeloid or  
 XX CC lymphoid cell disorders, platelet disorders such as thrombocytopenia and  
 XX CC for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
 XX CC growth, in tissue repair, healing of burns, incisions, ulcers, for  
 XX CC treating osteoporosis, osteoarthritis, bone degenerative disorders, and  
 XX CC periodontal disease. The polypeptides are also useful for gut protection  
 XX CC or regeneration and treatment of lung or liver fibrosis, reperfusion  
 XX CC injury in various tissues, various immune deficiencies and disorders  
 XX CC including severe combined immunodeficiency (SCID), bacterial or fungal  
 XX CC infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid

CC arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and  
 CC conditions, such as asthma or other respiratory problems. The  
 CC polypeptides are involved in thrombolysis or thrombosis and are useful in  
 CC treatment of various coagulation disorders (including hereditary  
 CC disorders such as haemophilia) or to enhance coagulation and other  
 CC haemostatic events in treating wounds resulting from trauma, surgery or  
 CC other causes. The polypeptides exhibit immune stimulating or immune  
 CC suppressing activity, and are useful for treating autoimmune diseases or  
 CC cancer. They also inhibit the growth, infection or function of infectious  
 CC agents such as bacteria, fungi, viruses, effect biohythms or circadian  
 CC cycles of rhythms, fertility of male or female subjects, metabolism,  
 CC catapoleism, and anabolism. ABO00414-ABO00749 represent the novel  
 CC polypeptides of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC XX

SQ Sequence 167 AA;

Query Match 56.5%; Score 52; DB 6; Length 167;  
 Best Local Similarity 53.8%; Pred. No. 4.1;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 WIMNQMHIDEKN 14  
 | : | | : | |  
 Db 45 WYMNQFVPEEN 57

Search completed: July 29, 2004, 11:34:54  
 Job time : 48.6667 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:32:25 ; Search time 13 Seconds

(Without alignments)  
59.568 Million cell updates/sec

Title: US-10-040-128-3

Sequence: 1 DWIMNQHIDEKNT 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCRTUS COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	108	4	US-09-187-859-6
2	92	100.0	108	4	US-09-839-542B-6
3	92	100.0	108	4	US-09-535-852-6
4	92	100.0	780	1	US-08-188-228-50
5	92	100.0	780	1	US-08-332-643-44
6	92	100.0	780	1	US-08-332-638-50
7	92	100.0	780	1	US-08-332-638-50
8	49	53.3	110	4	US-09-187-859-13
9	49	53.3	110	4	US-09-839-542B-13
10	49	53.3	532	1	US-08-188-228-44
11	49	53.3	532	1	US-08-332-638-44
12	49	53.3	793	1	US-08-188-228-54
13	49	53.3	793	1	US-08-332-643-48
14	49	53.3	793	1	US-08-332-638-54
15	49	53.3	799	1	US-08-188-228-42
16	49	53.3	799	1	US-08-332-638-42
17	45	48.9	106	4	US-09-187-859-44
18	45	48.9	106	4	US-09-187-859-45
19	45	48.9	106	4	US-09-305-927-1
20	45	48.9	106	4	US-09-305-927-2
21	45	48.9	106	4	US-09-305-927-3
22	45	48.9	106	4	US-09-073-040-12
23	45	48.9	106	4	US-09-073-040-13
24	45	48.9	106	4	US-09-839-542B-44
25	45	48.9	106	4	US-09-839-542B-45
26	45	48.9	106	4	US-09-264-516A-1
27	45	48.9	106	4	US-09-264-516A-2

28	45	48.9	106	4	US-09-264-516A-3	Sequence 3, Appli
29	45	48.9	106	4	US-09-234-395-4	Sequence 4, Appli
30	45	48.9	106	4	US-09-234-395-5	Sequence 5, Appli
31	45	48.9	106	4	US-09-305-928-4	Sequence 4, Appli
32	45	48.9	106	4	US-09-305-928-5	Sequence 5, Appli
33	45	48.9	110	4	US-09-187-859-2	Sequence 4, Appli
34	45	48.9	110	4	US-09-187-859-22	Sequence 22, Appli
35	45	48.9	110	4	US-09-839-542B-4	Sequence 4, Appli
36	45	48.9	110	4	US-09-839-542B-22	Sequence 22, Appli
37	45	48.9	110	4	US-09-535-852-4	Sequence 4, Appli
38	45	48.9	110	4	US-09-535-852-22	Sequence 22, Appli
39	45	48.9	615	2	US-08-738-349-12	Sequence 12, Appli
40	45	48.9	693	2	US-08-738-349-6	Sequence 6, Appli
41	45	48.9	796	1	US-08-188-228-58	Sequence 58, Appli
42	45	48.9	796	1	US-08-332-643-52	Sequence 52, Appli
43	45	48.9	796	1	US-08-332-638-58	Sequence 58, Appli
44	45	48.9	796	2	US-08-738-349-2	Sequence 2, Appli
45	45	48.9	796	2	US-08-738-349-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-09-187-859-6  
Sequence 6, Application US/09187859A  
Patent No. 6358920  
GENERAL INFORMATION:  
APPLICANT: Blaschuk, Orest W.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
FILE REFERENCE: 100086.407C1  
CURRENT APPLICATION NUMBER: US/09/187,859A  
CURRENT FILING DATE: 1998-11-06  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 108  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-187-859-6

Query Match 100.0%; Score 92; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMNQHIDEKNT 15  
DB 5 DWIMNQHIDEKNT 19

RESULT 2  
US-09-839-542B-6  
Sequence 6, Application US/09839542B  
Patent No. 6569996  
GENERAL INFORMATION:  
APPLICANT: Blaschuk, Orest W.  
APPLICANT: Symonds, James Matthew  
APPLICANT: Gour, Barbara J.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
FILE REFERENCE: 100086.407D1  
CURRENT APPLICATION NUMBER: US/09/839,542B  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 4052  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 108  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-839-542B-6

Query Match 100.0%; Score 92; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNT 15  
Db 5 DWIMQMHIDEKNT 19

## RESULT 3

US-09-535-852-6  
; Sequence 6, Application US/09535852  
; Patent No. 6638911  
; GENERAL INFORMATION:  
; APPLICANT: Blachuk, Orest W.  
; APPLICANT: Symonds, James M.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING  
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C6  
; CURRENT APPLICATION NUMBER: US/09/535,852  
; CURRENT FILING DATE: 2001-05-21  
; NUMBER OF SEQ ID NOS: 2009  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-535-852-6

Query Match 100.0%; Score 92; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNT 15  
Db 5 DWIMQMHIDEKNT 19

## RESULT 4

US-08-188-228-50  
; Sequence 50, Application US/08188228  
; Patent No. 5597725  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/188,228  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/049,460  
; FILING DATE: 19 APR 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,643  
; FILING DATE: 17 APR 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5597725and, Greta E.  
; REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31340  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-188-228-50

Query Match 100.0%; Score 92; DB 1; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNT 15  
Db 44 DWIMQMHIDEKNT 58

## RESULT 5

US-08-332-643-44  
; Sequence 44, Application US/08332643  
; Patent No. 5639634  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,643  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/872,643  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5639634and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-332-643-44

Query Match 100.0%; Score 92; DB 1; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNT 15  
|||||



Db 44 DWINOMHIDEKNT 58

## RESULT 6

US-08-332-638-50  
; Sequence 50, Application US/08332638  
; Patent No. 5646250  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,638  
; FILING DATE: 01-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,643  
; FILING DATE: 17 APR 1992  
; APPLICATION NUMBER: US/08/049,460  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5646250and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31340  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 780 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-332-638-50

Query Match 100.0%; Score 92; DB 1; Length 780;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINOMHIDEKNT 15  
Db 44 DWINOMHIDEKNT 58

## RESULT 7

US-09-187-859-13  
; Sequence 13, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; NUMBER OF SEQ ID NOS: 1998-11-06  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 13

; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-187-859-13

Query Match 53.3%; Score 49; DB 4; Length 110;  
Best Local Similarity 63.6%; Pred. No. 1.6;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WWINOMHIDE 12  
Db 6 WWINOMFVLEE 16

## RESULT 8

US-09-839-542B-13  
; Sequence 13, Application US/09839542B  
; Patent No. 656996  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-839-542B-13

Query Match 53.3%; Score 49; DB 4; Length 110;  
Best Local Similarity 63.6%; Pred. No. 1.6;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WWINOMHIDE 12  
Db 6 WWINOMFVLEE 16

## RESULT 9

US-09-535-852-13  
; Sequence 13, Application US/09535852  
; Patent No. 6638911  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James M.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING  
; FILE REFERENCE: 100086.407C6  
; CURRENT APPLICATION NUMBER: US/09/535,852  
; CURRENT FILING DATE: 2001-05-21  
; NUMBER OF SEQ ID NOS: 2009  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-535-852-13

Query Match 53.3%; Score 49; DB 4; Length 110;  
Best Local Similarity 63.6%; Pred. No. 1.6;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WWINOMHIDE 12  
Db 6 WWINOMFVLEE 16

RESULT 10  
US-08-188-228-44  
Sequence 44, Application US/08188228  
Patent No. 5597725  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,228  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,460  
FILING DATE: 19 APR 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5597725and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-228-44

Query Match 53.3%; Score 49; DB 1; Length 532;  
Best Local Similarity 63.6%; Pred. No. 8.9;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12  
DB 63 WYNNOMFVLEE 73

RESULT 11  
US-08-332-638-44  
Sequence 44, Application US/08332638  
Patent No. 5646250  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,638  
FILING DATE: 01-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
APPLICATION NUMBER: US/08/049,460  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5646250and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-332-638-44

Query Match 53.3%; Score 49; DB 1; Length 532;  
Best Local Similarity 63.6%; Pred. No. 8.9;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12  
DB 63 WYNNOMFVLEE 73

RESULT 12  
US-08-188-228-54  
Sequence 54, Application US/08188228  
Patent No. 5597725  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,228  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,460  
FILING DATE: 19 APR 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5597725and, Greta E.  
REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-228-54

Query Match 53.3%; Score 49; DB 1; Length 793;  
Best Local Similarity 63.6%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12  
|:|||||:|  
56 WWINOMFVLEE 66

RESULT 13  
US-08-332-643-48  
Sequence 48, Application US/08332643  
Patent No. 5639634  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,643  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/872,643  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5639634and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30795  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-332-643-48

Query Match 53.3%; Score 49; DB 1; Length 793;  
Best Local Similarity 63.6%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 WINNOMHIDE 12  
|:|||||:|

Db 56 WWINOMFVLEE 66

RESULT 14  
US-08-332-638-54  
Sequence 54, Application US/08332638  
Patent No. 5646250  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,638  
FILING DATE: 01-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
APPLICATION NUMBER: US/08/049,460  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5646250and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-332-638-54

Query Match 53.3%; Score 49; DB 1; Length 793;  
Best Local Similarity 63.6%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12  
|:|||||:|  
56 WWINOMFVLEE 66

RESULT 15  
US-08-188-228-42  
Sequence 42, Application US/08188228  
Patent No. 5597725  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA

ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,228  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,460  
FILING DATE: 19 APR 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/872,643  
FILING DATE: 17 APR 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 559725and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 799 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-228-42

Query Match 53.3%; Score 49; DB 1; Length 799;  
Best Local Similarity 63.6%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 MINOMHIDE 12  
|:|:|:|:|:  
Db 63 WWMNQMFVLEE 73

Search completed: July 29, 2004, 11:38:39  
Job time : 13 secs



CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas, The  
 CC adenocarcinomas, adenomas, leukaemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vasculature in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
 CC domain 1 antibody binding peptide

SO Sequence 15 AA;

Query Match 100.0%; Score 92; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.3e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DWIMNQHIDEKNT 15  
 DB 1 DWIMNQHIDEKNT 15

RESULT 2

AA08785 standard; protein; 104 AA.

AA08785;

27-FEB-2002 (first entry)

Human VE-cadherin N-terminal domain 1 polypeptide.

XX VE-cadherin; angiogenesis; neoplastic disease;  
 XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
 XX paracellular permeability; solid tumour; collagenous vascular disease;  
 XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
 XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;  
 XX carcinoma; sarcoma; adenocarcinoma; adenoma; gene therapy;  
 XX leukaemic tumour; lymphoid tumours; cell proliferation disorder;  
 XX vascularisation; blood vessel proliferation disorder; fibrotic disorder;  
 XX age-related muscular degeneration; tumour neovascularisation; human;  
 XX VE-cadherin N-terminal domain 1.

OS Homo sapiens.

PN WO200175109-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US010505.

PR 31-MAR-2000; 2000US-00540967.

PA (IMCL-) IMCLONE SYSTEMS INC.

PI Liao F, Hicklin DJ, Bohlen P;

DR WPI; 2001-656988/75.

XX New antibody antagonists of VE-cadherin, which does not adversely affect  
 XX vascular permeability, useful for inhibiting angiogenesis or tumour  
 XX metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.

XX Example 2; Fig 2; 4pp; English.

XX The method relates to an antibody or an antibody fragment, capable of

CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 CC mediated adherens junction formation in vitro, but does not exert any  
 CC significant or substantial effect on paracellular permeability in vitro.  
 CC The antibody or antibody fragment is capable of specifically binding to a  
 CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma) or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas, The  
 CC adenocarcinomas, adenomas, leukaemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vasculature in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a human VE-cadherin N-  
 CC terminal domain 1 polypeptide

SO Sequence 104 AA;

Query Match 100.0%; Score 92; DB 4; Length 104;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DWIMNQHIDEKNT 15  
 DB 1 DWIMNQHIDEKNT 15

RESULT 3

AA64578 standard; peptide; 108 AA.

AA64578;

02-MAR-2000 (first entry)

Nonclassical cadherin extracellular domain SEQ ID NO:6.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 XX inhibition; cadherin extracellular domain; cell adhesion recognition;  
 XX OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 XX cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 XX cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 XX desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 XX rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 XX neurological disease.

OS Mammalia.

PN WO9957149-A2.

PD 11-NOV-1999.

PF 05-MAY-1999; 99WO-CA000363.

PR 05-MAY-1998; 98US-00073040.

PR 06-NOV-1998; 98US-0018785.

PR 20-JAN-1999; 99US-00234395.

PR 08-MAR-1999; 99US-00264516.

PA (ADHE-) ADHEREX TECHNOLOGIES INC.

PI Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.

XX New cadherin modulating agents, used for modulating nonclassical cadherin  
 XX mediated functions for treating e.g. cancers, obesity, rheumatoid

PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
XX  
PS Disclosure; Fig 2; 252pp; English.  
XX  
CC The present invention describes cadherin modulating agents (MA)  
CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
CC recognition (CAR) sequence. The MAs can be used for modulating  
CC nonclassical cadherin-mediated functions. They can be used for e.g.  
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
CC mammal, enhancing delivery of a drug through the skin of a mammal,  
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
CC expressing cell, preventing or treating obesity in a mammal, stimulating  
CC blood vessel regression in a mammal, enhancing drug delivery to the  
CC central nervous system, treating a demyelinating neurological disease,  
CC increasing vasopermeability in a mammal, enhancing adhesion of  
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
CC -related macular degeneration, multiple sclerosis and diabetes. The  
CC products can also be used for detection and diagnosis and in bioreactors.  
CC AAY64572 to AAY64572 represent specifically claimed peptides, and  
CC AAY64573 to AAY64643 and AAZ31183 to AAZ31186 represent sequences used in  
CC the exemplification of the present invention  
XX  
SQ Sequence 108 AA;  
  
Query Match 100.0%; Score 92; DB 3; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DWTWQMHIIDEKNT 15  
| | | | | | | | | | | | | | | | | | | | | |  
DB 5 DWTWQMHIIDEKNT 19  
  
RESULT 4  
AAU11276  
ID AAU11276 standard; protein; 566 AA.  
XX  
AC AAU11276;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human recombinant KVE702 polypeptide.  
XX  
KW Human; KV protein; KVE702; angiogenic factor antagonist; angiogenesis;  
KW vascular endothelial structure regulator; cancer; vascular endothelium;  
KW tumour; blood vessel; epithelial cell; cytosstatic; gene therapy.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
WO200183562-A2.  
XX  
PD 08-NOV-2001.  
XX  
PF 04-MAY-2001; 2001WO-GB001956.  
XX  
PR 04-MAY-2000; 2000GB-00010630.  
XX  
PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.  
XX  
PI Jiang WG;  
XX  
DR MPI; 2002-066521/09.  
DR N-PSDB; AAS16965.  
PT Novel nucleic acid for inhibition of angiogenesis and prophylaxis or  
PT treatment of cancer, comprises both an angiogenic factor antagonist and a

PT vascular endothelial structure regulator.  
XX  
PS Disclosure; Fig 2; 66pp; English.  
XX  
CC The invention relates to recombinant nucleic acid sequences that encode  
CC both an angiogenic factor antagonist and a vascular endothelial structure  
CC regulator, collectively referred to as a KV protein. KV DNA and proteins,  
CC and also cells and compositions containing the sequences, are useful in  
CC treating or preventing cancer or angiogenesis in mammals, in particular  
CC humans. KV proteins are also useful for regulation of the development of  
CC blood vessels and their formation, in the vascular endothelium and/or  
CC tumour. The KVE702 gene and its fragments are useful in transfection of  
CC human epithelial cells and to generate products suitable for angiogenesis  
CC intervention. This sequence represents the human KVE702 protein  
XX  
SQ Sequence 566 AA;  
  
Query Match 100.0%; Score 92; DB 5; Length 566;  
Best Local Similarity 100.0%; Pred. No. 7.2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 DWTWQMHIIDEKNT 15  
| | | | | | | | | | | | | | | | | | | | | |  
DB 474 DWTWQMHIIDEKNT 488  
  
RESULT 5  
AAU11279  
ID AAU11279 standard; protein; 594 AA.  
XX  
AC AAU11279;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human recombinant J37 polypeptide.  
XX  
KW Human; KV protein; KVE702; angiogenic factor antagonist; angiogenesis;  
KW vascular endothelial structure regulator; cancer; vascular endothelium;  
KW tumour; blood vessel; epithelial cell; cytosstatic; gene therapy; J37.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FH FT Misc-difference 518  
FT FT /note= "Encoded by ATC"  
XX  
PN WO200183562-A2.  
XX  
PD 08-NOV-2001.  
XX  
PF 04-MAY-2001; 2001WO-GB001956.  
XX  
PR 04-MAY-2000; 2000GB-00010630.  
XX  
PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.  
XX  
PI Jiang WG;  
XX  
DR MPI; 2002-066521/09.  
DR N-PSDB; AAS16975.  
PT Novel nucleic acid for inhibition of angiogenesis and prophylaxis or  
PT treatment of cancer, comprises both an angiogenic factor antagonist and a  
PT vascular endothelial structure regulator.  
XX  
PS Disclosure; Fig 15; 66pp; English.  
XX  
CC The invention relates to recombinant nucleic acid sequences that encode  
CC both an angiogenic factor antagonist and a vascular endothelial structure  
CC regulator, collectively referred to as a KV protein. KV DNA and proteins,  
CC and also cells and compositions containing the sequences, are useful in  
CC treating or preventing cancer or angiogenesis in mammals, in particular

CC humans. KV proteins are also useful for regulation of the development of  
 CC blood vessels and their formation, in the vascular endothelium and/or  
 CC tumour. The KVE702 gene and its fragments are useful in transfection of  
 CC human epithelial cells and to generate products suitable for angiogenesis  
 CC intervention. This sequence represents the human J37 protein used in  
 CC methods of the invention

XX Sequence 594 AA;

Query Match 100.0%; Score 92; DB 5; Length 594;

Best Local Similarity 100.0%; Pred. No. 7.6e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWTWQMHIDEKNT 15  
 |||||||

DB 49 DWTWQMHIDEKNT 63

#### RESULT 6

AAW25634  
 ID AAW25634 standard; protein; 780 AA.

XX AAW25634;

DT 25-MAR-2003 (revised)  
 DT 03-NOV-1997 (first entry)

DE Human cadherin-5.

XX Human; cadherin; rat; calcium-dependent cell adhesion protein;  
 KW superfamily; cytoskeleton; eatenin; cancer.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 622  
 FT /note= "encoded by AAG"

XX US5646250-A.

XX 08-JUL-1997.

XX 01-NOV-1994; 94US-00332638.

XX 17-APR-1992; 92US-00872643.

XX 19-APR-1993; 93US-00049460.

XX (DOHE-) DOHENY EYE INST.

XX Suzuki S;

XX WPI; 1997-362997/33.

XX N-PSDB; AAT85401.

XX Human and rat cadherin polypeptide(s) - mediate cell-cell adhesion.

XX Claim 1; Col 69-72; 56pp; English.

XX This sequence represents human cadherin-5. The invention specifically  
 CC provides details of human cadherin-5, -8, -11, -12 and -13, and rat  
 CC cadherin-8, -5, -11 and -13. Cadherins are calcium-dependent cell  
 CC adhesion proteins. They are glycosylated integral membrane proteins that  
 CC have an N-terminal extracellular domain that determines binding  
 CC specificity, a hydrophobic membrane spanning region and a C-terminal  
 CC cytoplasmic domain, which is highly conserved among members of the  
 CC superfamily. The C-terminal domain interacts with the cytoskeleton  
 CC through eatenins and other cytoskeleton-associated proteins. The novel  
 CC cadherin proteins may be used in the analysis of the role of cadherins in  
 CC various cancers. Sequence analysis of the cadherin proteins also allows  
 CC investigation of the structure and function of cadherin. The cadherin  
 CC proteins may be isolated by using anti-cadherin antibodies. These  
 CC antibodies may also be used to modulate the activity of cadherin and to  
 CC determine the tissue specific distribution of cadherin proteins. Each

CC subclass of cadherins has a unique tissue distribution pattern. (Updated  
 CC on 25-MAR-2003 to correct PF field.)

XX Sequence 780 AA;

QY Query Match 100.0%; Score 92; DB 2; Length 780;

Best Local Similarity 100.0%; Pred. No. 1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWTWQMHIDEKNT 15  
 |||||||

DB 44 DWTWQMHIDEKNT 58

#### RESULT 7

AAW13130  
 ID AAW13130 standard; protein; 780 AA.

XX AAW13130;

DT 25-MAR-2003 (revised)  
 DT 13-MAY-1997 (first entry)

DE Full length human cadherin-5.

XX Ca2+ dependent; cell adhesion protein; placental; cadherin; rat; brain;  
 KW human; antibody; purification; determination; tissue expression;  
 KW binding antagonist; calcium ion.

XX Homo sapiens.

XX US5597725-A.

XX 28-JAN-1997.

XX 26-JAN-1994; 94US-00188228.

XX 17-APR-1992; 92US-00872643.

XX 19-APR-1993; 93US-00049460.

XX (DOHE-) DOHENY EYE INST.

XX Suzuki S;

XX WPI; 1997-108328/10.

XX N-PSDB; AAT61921.

XX Antibodies to cadherin proteins - useful as cadherin antagonists, etc.

XX Claim 5; Col 75-78; 59pp; English.

XX The present sequence is full length human cadherin-5, which is a Ca2+  
 CC dependent cell adhesion protein. The human cadherin cDNA was isolated  
 CC from a placental cDNA library, using probes based on homologous rat  
 CC cadherin cDNA. Antibodies or fragments that specifically bind the human  
 CC cadherin can be used to purify the cadherin, determine its tissue  
 CC expression and antagonise its ligand/antiligand binding activities.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 780 AA;

Query Match 100.0%; Score 92; DB 2; Length 780;

Best Local Similarity 100.0%; Pred. No. 1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWTWQMHIDEKNT 15  
 |||||||

DB 44 DWTWQMHIDEKNT 58

#### RESULT 8

ABB81477  
 ID ABB81477 standard; protein; 784 AA.



[illegible]

```

XX      21-JAN-2003   (first entry)
DT
XX      Angiogenesis-associated human protein sequence #52.
DE
XX      Human; angiogenesis-associated transcript; angiogenesis;
KM      angiogenesis-associated disease; cancer; cytostatic.
XX
OS      Homo sapiens.
PN      WO200279492-A2.
PD      10-OCT-2002.
PF      14-FEB-2002; 2002WO-US004915.
XX
PR      14-FEB-2001; 2001US-00784356.
PR      22-FEB-2001; 2001US-00791390.
PR      19-APR-2001; 2001US-0285475P.
PR      03-AUG-2001; 2001US-0310025P.
PR      13-NOV-2001; 0350666P.
PR      29-NOV-2001; 2001US-0334244P.
XX
PA      (EOSB-) EOS BIOTECHNOLOGY INC.
PI
XX      Murray R, Glynn R, Watson SR, Aziz N;
DR      WPI; 2003-040681/03.
DR      N-PDSB; ABX08791.
PT
XX      Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT      treating cancer by contacting a sample with a polynucleotide that
PT      exhibits changes in expression level as a function of time in tissue
XX      undergoing angiogenesis.
PS
XX      Example 2; Page 230; 291pp; English.
CC
XX      The present invention relates to methods and compositions for detecting
CC      an angiogenesis-associated transcript in a cell in a patient. The method
CC      involves contacting a biological sample from the patient with a
CC      polynucleotide that selectively hybridises to a sequence at least 80%
CC      identical to any of the angiogenesis-associated human polynucleotide
CC      sequences given in the specification. These angiogenesis-associated
CC      polynucleotide sequences comprise genes that exhibit changes in
CC      expression levels as a function of time in tissue undergoing
CC      angiogenesis. The method and the polynucleotide sequences of the
CC      invention are useful for diagnosing and treating angiogenesis and
CC      angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC      sequences are also useful in the gene therapy of such disorders. The
CC      angiogenesis-associated proteins encoded by the polynucleotide sequences
CC      are useful as a vaccine for therapeutic and prophylactic immunisation.
CC      ABU03456-ABU03569 represent angiogenesis-associated protein sequences.
XX
XX      Sequence 784 AA:
SQ
Query Match          100.0%; Score 92; DB 6; Length 784;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY      1 DWINQMHIIDEKXT 15
        |||
DB      48 DWINQMHIIDEKXT 62
RESULT 10
ABU56570
ID      ABU56570 standard; protein; 784 AA.
XX
AC      ABU56570;
XX
DT      02-APR-2003   (first entry)
XX      Lung cancer-associated polypeptide #163.
DE
```

XX lung cancer-associated polypeptide; cytosratic; emphysema;  
XX antiinflammatory; antiastrmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; actuma; bronchiectasis.  
XX  
OS unidentified.  
XX  
PN WO200286443-A2.  
XX  
PD 31-OCT-2002.  
XX  
PF 18-APR-2002; 2002WO-US012476.  
XX  
PR 18-APR-2001; 2001US-0284770P.  
PR 10-MAY-2001; 2001US-0290492P.  
PR 09-NOV-2001; 2001US-0339245P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334370P.  
PR 12-APR-2002; 2002US-0372246P.  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
PI Aziz N, Murray R;  
DR WPI; 2003-093161/08.  
XX  
DR N-PsDB; ABX76299.  
XX  
PT Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the  
PT patient with a polynucleotide that exhibits increased or decreased  
PS expression in lung cancer.  
XX  
PS Claim 27; Page 314; 453pp; English.

CC The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung cancer-  
CC associated polynucleotides and polypeptides are used for identifying a  
CC compound that modulates a lung cancer-associated polypeptide, for  
CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
CC cancer in a patient and for treating a mammal having lung cancer by  
CC administering a modulatory compound identified. The methods are useful  
CC for treating lung cancer, such as small cell lung cancer, non-small cell  
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, and  
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
CC for diagnostic purposes and as targets for screening for therapeutic  
CC compounds that modulate lung cancer, such as antibodies. Sequences  
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the  
CC invention  
XX  
XX

SO Sequence 784 AA;

Query Match 100.0%; Score 92; DB 6; Length 784;  
Best Local Similarity 100.0%; Pred No. le-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DWIWMQHIDEKNT 15  
| | | | | | | | | | | | | | | | |  
Db 48 DWIWMQHIDEKNT 62

RESULT 11  
ADA03363  
ID ADA03363 standard; protein; 784 AA.  
AC ADA03363;  
XX

DT	06-NOV-2003	(first entry)	
DE		Human vascular endothelial cadherin protein.	
XX			
XX		cytostatic; antidiabetic; ophthalmological; antiarthritic;	
XX		endostatin receptor; angiogenesis inhibitor; human;	
KW		vascular endothelial cadherin; angiogenesis; blood vessel; cancer;	
KW		diabetic retinopathy; rheumatoid arthritis; receptor.	
OS		Homo sapiens.	
XX			
PN		WO2003038096-A1.	
XX			
PD		08-MAY-2003.	
XX			
PF		28-OCT-2002; 2002WO-JP011143.	
XX			
PR		30-OCT-2001; 2001JP-00333295.	
XX			
PA		(KUBO/) KUBOTA S.	
XX		(NEMO/) NEMOTO T.	
XX			
PI		Kubota S, Nemoto T;	
XX			
DR		WPI, 2003-430528/40.	
XX		N-PSDB; ADA03362.	
PT		Vascular endothelial cadherin is a receptor to endostatin for screening	
PT		endostatin agonists and antagonists as angiogenesis inhibitors and	
PT		remedies for cancer development and proliferation.	
XX			
PS		Claim 2; Page 44-47; 56pp; Japanese.	
XX			
CC		The invention relates to a DNA encoding an endostatin receptor protein	
CC		(human vascular endothelial (VE) cadherin (I)) or its partial peptides or	
CC		encoding a protein derived from (I) by addition, deletion or substitution	
CC		of one or more amino acid residues and having similar binding activity to	
CC		endostatin. The sequence can be used in the regulation of angiogenesis in	
CC		the treatment and prevention of diseases associated with the formation of	
CC		new blood vessels, including cancer, cancer proliferation, diabetic	
CC		retinopathy and rheumatoid arthritis. This sequence corresponds to the VE	
CC		cadherin protein which serves as an endostatin receptor protein.	
XX			
SO		Sequence 784 AA;	
	Query Match	100.0%;	Score 92; DB 6; Length 784;
	Best Local Similarity	100.0%;	Pred. No. 1e-05;
	Matches 15; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 DWINQMHIIDEKNT 15		
	48 DWINQMHIIDEKNT 62		
RESULT 12			
ID	AAU08778		
AC	AAU08778 standard; peptide; 15 AA.		
XX			
DT	27-FEB-2002	(first entry)	
XX			
DE		VE-cadherin N-terminal domain 1 antibody binding peptide #2.	
KW		VE-cadherin; antibody binding peptide; angiogenesis; neoplastic disease;	
KW		VE-cadherin mediated adherens junction formation; autoimmune disease;	
KW		paracellular permeability; solid tumour; collagenous vascular disease;	
KW		rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;	
KW		retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;	
KW		carcinoma; sarcoma; adenocarcinoma; adenocarcinoma; adenoma; gene therapy;	
KW		leukaemic tumour; lymphoid tumours; cell proliferation disorder;	
KW		vascularisation; blood vessel proliferation disorder; fibrotic disorder;	
KW		age-related muscular degeneration; tumour neovascularisation;	

KW VE-cadherin N-terminal domain 1.  
 XX Mammalia.  
 OS  
 XX WO200175109-A2.  
 XX  
 XX PD 11-OCT-2001.  
 XX  
 XX PF 30-MAR-2001; 2001WO-US010505.  
 XX  
 XX PR 31-MAR-2000; 2000US-00540967.  
 XX  
 XX PA (IMCL-) IMCLONE SYSTEMS INC.  
 XX  
 XX PI Liao F, Hicklin DJ, Bohlen P;  
 XX WPL; 2001-656988/75.  
 XX  
 XX PT New antibody antagonists of VE-cadherin, which does not adversely affect  
 PT vascular permeability, useful for inhibiting angiogenesis or tumor  
 PT metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.  
 XX  
 XX PS Claim 1; Page 32; 44pp; English.  
 XX  
 CC The method relates to an antibody or an antibody fragment, capable of  
 CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 CC mediated adherens junction formation in vitro, but does not exert any  
 CC significant or substantial effect on paracellular permeability in vitro.  
 CC The antibody or antibody fragment is capable of specifically binding to  
 CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma), or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 CC adenocarcinomas, adenomas, leukemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vasculature in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a VE-cadherin N-terminal  
 CC domain 1 antibody binding peptide  
 CC  
 CC SQ Sequence 15 AA;  
 CC  
 CC Query Match 94.6%; Score 87; DB 4; Length 15;  
 CC Best Local Similarity 100.0%; Pred. No. 7.7e-07;  
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC Oy 1 DWINQWHIDEKN 14  
 CC |||||  
 CC Db 1 DWINQWHIDEKN 14  
 CC  
 CC RESULT 13  
 CC AAU08786  
 CC ID AAU08786 standard; protein; 103 AA.  
 CC AC  
 CC XX AAU08786;  
 CC XX  
 CC XX 27-FEB-2002 (first entry)  
 CC XX  
 CC DE Murine VE-cadherin N-terminal domain 1 polypeptide.  
 CC  
 CC XX VE-cadherin; angiogenesis; neoplastic disease;  
 CC XX VE-cadherin mediated adherens junction formation; autoimmune disease;  
 CC XX paracellular permeability; solid tumour; collagenous vascular disease;  
 CC XX rheumatoid arthritis; ophthalmological condition; diabetic retinopathy;  
 CC XX retrolental fibroplasia; neovascular glaucoma; tumour metastasis; glioma;

KW carcinoma; sarcoma; adenocarcinoma; adenoma; gene therapy;  
 KW leukemic tumour; lymphoid tumours; cell proliferation disorder;  
 KW vascularisation; blood vessel proliferation disorder; fibrotic disorder;  
 KW age-related muscular degeneration; tumour neovascularisation; mouse;  
 KW VE-cadherin N-terminal domain 1.  
 XX  
 XX OS Mus sp.  
 XX  
 XX XX WO200175109-A2.  
 XX  
 XX PD 11-OCT-2001.  
 XX  
 XX PF 30-MAR-2001; 2001WO-US010505.  
 XX  
 XX PR 31-MAR-2000; 2000US-00540967.  
 XX  
 XX PA (IMCL-) IMCLONE SYSTEMS INC.  
 XX  
 XX PI Liao F, Hicklin DJ, Bohlen P;  
 XX WPL; 2001-656988/75.  
 XX  
 XX PT New antibody antagonists of VE-cadherin, which does not adversely affect  
 PT vascular permeability, useful for inhibiting angiogenesis or tumor  
 PT metastasis, e.g. autoimmune disease, carcinomas or leukemic tumors.  
 XX  
 XX PS Example 2; Fig 2; 44pp; English.  
 XX  
 CC The method relates to an antibody or an antibody fragment, capable of  
 CC specifically binding to a VE-cadherin, and of inhibiting VE-cadherin  
 CC mediated adherens junction formation in vitro, but does not exert any  
 CC significant or substantial effect on paracellular permeability in vitro.  
 CC The antibody or antibody fragment is capable of specifically binding to a  
 CC site on a VE-cadherin within the 15 N-terminal amino acids of domain 1 of  
 CC a VE-cadherin. The antibody or antibody fragment is useful for inhibiting  
 CC angiogenesis (e.g. angiogenesis that is associated with a neoplastic  
 CC disease, a solid tumour, an autoimmune disease, collagenous vascular  
 CC disease, rheumatoid arthritis, an ophthalmological condition, diabetic  
 CC retinopathy, retrolental fibroplasia or neovascular glaucoma), or tumour  
 CC metastasis (e.g. carcinomas, gliomas, sarcomas, adenocarcinomas,  
 CC adenocarcinomas, adenomas, leukemic tumours or lymphoid tumours). The  
 CC composition or antibody is also useful for treating a cell proliferation  
 CC disorder associated with vascularisation (e.g. blood vessel proliferation  
 CC disorders, fibrotic disorders, angiogenesis, tumour growth, tumour  
 CC metastasis, rheumatoid arthritis or age-related muscular degeneration).  
 CC These may also be used for reducing or inhibiting tumour vasculature in a  
 CC mammal. The nucleic acid that encodes the antibody or is useful in gene  
 CC therapy, particularly for inhibiting angiogenesis or tumour  
 CC neovascularisation. This sequence represents a murine VE-cadherin N-  
 CC terminal domain 1 polypeptide  
 CC  
 CC SQ Sequence 103 AA;  
 CC  
 CC Query Match 94.6%; Score 87; DB 4; Length 103;  
 CC Best Local Similarity 100.0%; Pred. No. 6.5e-06;  
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC Oy 1 DWINQWHIDEKN 14  
 CC |||||  
 CC Db 1 DWINQWHIDEKN 14  
 CC  
 CC RESULT 14  
 CC AAU08777  
 CC ID AAU08777 standard; peptide; 15 AA.  
 CC AC  
 CC XX AAU08777;  
 CC XX  
 CC XX 27-FEB-2002 (first entry)  
 CC XX  
 CC DE VE-cadherin N-terminal domain 1 antibody binding peptide #1.  
 CC  
 CC XX VE-cadherin; antibody binding peptide; angiogenesis; neoplastic disease;



CC arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and  
 CC conditions, such as asthma or other respiratory problems. The  
 CC polypeptides are involved in thrombolysis or thrombosis and are useful in  
 CC treatment of various coagulation disorders (including hereditary  
 CC disorders such as haemophilia) or to enhance coagulation and other  
 CC haemostatic events in treating wounds resulting from trauma, surgery or  
 CC other causes. The polypeptides exhibit immune stimulating or immune  
 CC suppressing activity, and are useful for treating autoimmune diseases or  
 CC cancer. They also inhibit the growth, infection or function of infectious  
 CC agents such as bacteria, fungi, viruses, effect biorhythms or circadian  
 CC cycles of rhythms, fertility of male or female subjects, metabolism,  
 CC catabolism, and anabolism. AB000414-AB000749 represent the novel  
 CC polypeptides of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 167 AA;

Query Match 62.0%; Score 57; DB 6; Length 167;  
 Best Local Similarity 57.1%; Pred. No. 0.54;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 MWNQMHIDEKNT 15  
 | : | | | : | | | |  
 Db 45 MWNQPFVPEHMT 58

Search completed: July 29, 2004, 11:34:54  
 Job time : 46.6667 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:37:16 ; Search time 38.333 Seconds

(without alignments)  
122.745 Million cell updates/sec

Title: US-10-040-128-3

Perfect score: 92  
Sequence: 1 DWIMQMHIDEKNT 15

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10D\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	15	US-10-040-128-3	Sequence 3, Appli
2	92	100.0	108	US-10-006-869-6	Sequence 6, Appli
3	92	100.0	108	US-10-395-032-6	Sequence 6, Appli
4	92	100.0	784	US-09-905-983-52	Sequence 52, Appli
5	92	100.0	784	US-10-211-462-105	Sequence 105, Appli
6	92	100.0	784	US-10-021-660-99	Sequence 99, Appli
7	92	100.0	784	US-10-174-677-23	Sequence 23, Appli
8	87	94.6	15	US-10-040-128-1	Sequence 1, Appli
9	73	79.3	15	US-10-040-128-1	Sequence 1, Appli
10	57	62.0	167	US-10-243-552-596	Sequence 596, Appli
11	57	62.0	772	US-09-978-295A-264	Sequence 264, App
12	57	62.0	772	US-09-978-687-264	Sequence 264, App
13	57	62.0	772	US-09-978-193A-264	Sequence 264, App
14	57	62.0	772	US-09-978-188A-264	Sequence 264, App
15	57	62.0	772	US-09-999-832A-264	Sequence 264, App

16	57	62.0	772	10	US-09-978-608A-264	Sequence 264, App
17	57	62.0	772	10	US-09-978-585A-264	Sequence 264, App
18	57	62.0	772	10	US-09-978-191A-264	Sequence 264, App
19	57	62.0	772	10	US-09-978-403A-264	Sequence 264, App
20	57	62.0	772	10	US-09-978-564A-264	Sequence 264, App
21	57	62.0	772	10	US-09-999-833A-264	Sequence 264, App
22	57	62.0	772	10	US-09-981-915A-264	Sequence 264, App
23	57	62.0	772	10	US-09-978-824-264	Sequence 264, App
24	57	62.0	772	10	US-09-918-585A-264	Sequence 264, App
25	57	62.0	772	10	US-09-978-423A-264	Sequence 264, App
26	57	62.0	772	10	US-09-978-193A-264	Sequence 264, App
27	57	62.0	772	10	US-09-999-830A-264	Sequence 264, App
28	57	62.0	772	10	US-09-978-757A-264	Sequence 264, App
29	57	62.0	772	10	US-09-978-187B-264	Sequence 264, App
30	57	62.0	772	10	US-09-978-643A-264	Sequence 264, App
31	57	62.0	772	10	US-09-978-375A-264	Sequence 264, App
32	57	62.0	772	10	US-09-978-299A-264	Sequence 264, App
33	57	62.0	772	10	US-09-978-188A-264	Sequence 264, App
34	57	62.0	772	10	US-09-978-194A-264	Sequence 264, App
35	57	62.0	772	10	US-09-999-829A-264	Sequence 264, App
36	57	62.0	772	10	US-09-978-194A-264	Sequence 264, App
37	57	62.0	772	10	US-09-978-299A-264	Sequence 264, App
38	57	62.0	772	10	US-09-978-544A-264	Sequence 264, App
39	57	62.0	772	10	US-09-978-665A-264	Sequence 264, App
40	57	62.0	772	10	US-09-978-802A-264	Sequence 264, App
41	57	62.0	772	12	US-10-164-749A-264	Sequence 264, App
42	57	62.0	772	12	US-10-206-915-106	Sequence 106, App
43	57	62.0	772	12	US-10-199-670-106	Sequence 106, App
44	57	62.0	772	12	US-10-201-858-106	Sequence 106, App
45	57	62.0	772	12	US-09-999-831A-264	Sequence 264, App

#### ALIGNMENTS

RESULT 1  
US-10-040-128-3  
Sequence 3, Application US/10040128  
Publication No. US20020160003A1  
GENERAL INFORMATION:  
APPLICANT: Ilaeo, Fang  
APPLICANT: Hicklin, Daniel J.  
TITLE OF INVENTION: Antibody Antagonists of VE-Cadherin Without Adverse Effects on  
TITLE OF INVENTION: Vascular Permeability  
FILE REFERENCE: 11245/46976  
CURRENT APPLICATION NUMBER: US/10/040,128  
CURRENT FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: WordPerfect 8.0 for Windows  
SEQ ID NO 3  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic peptide  
US-10-040-128-3

Query Match 100.0%; Score 92; DB 13; Length 15;  
Best Local Similarity 100.0%; Pred. No. 8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNT 15  
DB 1 DWIMQMHIDEKNT 15

RESULT 2  
US-10-006-869-6  
Sequence 6, Application US/10006869  
Publication No. US2003008216A1  
GENERAL INFORMATION:  
APPLICANT: Blaschuk, Orest W.

```

; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-869-6

Query Match          100.0%; Score 92; DB 14; Length 108;
Best Local Similarity 100.0%; Pred. No. 5.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DWIMNQHIDEKNT 15
DB      5 DWIMNQHIDEKNT 19

RESULT 3
US-10-395-032-6
; Sequence 6, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-395-032-6

Query Match          100.0%; Score 92; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 5.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DWIMNQHIDEKNT 15
DB      5 DWIMNQHIDEKNT 19

RESULT 4
US-09-905-983-52
; Sequence 52, Application US/09905983
; Patent No. US20020045591A1
; GENERAL INFORMATION:
; APPLICANT: Geiger, Benjamin
; APPLICANT: Ben-Ze'ev, Avri
; APPLICANT: Sadot, Einar
; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER
; FILE REFERENCE: 01/22326
; CURRENT APPLICATION NUMBER: US/09/905,983
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 52
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-983-52
```

```

Query Match          100.0%; Score 92; DB 9; Length 784;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DWIMNQHIDEKNT 15
DB      48 DWIMNQHIDEKNT 62

RESULT 5
US-10-211-462-105
; Sequence 105, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasna
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-105

Query Match          100.0%; Score 92; DB 12; Length 784;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DWIMNQHIDEKNT 15
DB      48 DWIMNQHIDEKNT 62

RESULT 6
US-10-021-660-99
; Sequence 99, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 784
```



;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-021-660-99

Query Match  
Best Local Similarity 100.0%; Score 92; DB 14; Length 784;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINOMHIDEKNT 15  
|||||  
Db 48 DWINOMHIDEKNT 62

RESULT 7  
US-10-174-677-23

; Sequence 23; Application US/10174677  
; Publication No. US20030190704A1  
; GENERAL INFORMATION:

; APPLICANT: Xie, Ting

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR

; FILE REFERENCE: 40716 (ID-012)  
; CURRENT APPLICATION NUMBER: US/10/174,677

; CURRENT FILING DATE: 2002-06-19  
; NUMBER OF SEQ ID NOS: 117

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23

; LENGTH: 784  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-174-677-23

Query Match  
Best Local Similarity 100.0%; Score 92; DB 14; Length 784;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINOMHIDEKNT 15  
|||||  
Db 48 DWINOMHIDEKNT 62

RESULT 8  
US-10-040-128-2

; Sequence 2; Application US/10040128  
; Publication No. US20020160003A1  
; GENERAL INFORMATION:

; APPLICANT: Liao, Fang

; APPLICANT: Hicklin, Daniel J.

; TITLE OF INVENTION: Antibody Antagonists of VE-Cadherin Without Adverse Effects on

; FILE REFERENCE: 11245/46976  
; CURRENT APPLICATION NUMBER: US/10/040,128

; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 2

; LENGTH: 15  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: synthetic peptide  
US-10-040-128-2

Query Match  
Best Local Similarity 94.6%; Score 87; DB 13; Length 15;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWINOMHIDEKNT 14  
|||||  
Db 1 DWINOMHIDEKNT 14

RESULT 9

US-10-040-128-1  
; Sequence 1; Application US/10040128  
; Publication No. US20020160003A1  
; GENERAL INFORMATION:

; APPLICANT: Liao, Fang

; APPLICANT: Hicklin, Daniel J.

; TITLE OF INVENTION: Antibody Antagonists of VE-Cadherin Without Adverse Effects on

; FILE REFERENCE: 11245/46976  
; CURRENT APPLICATION NUMBER: US/10/040,128

; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 1

; LENGTH: 15  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: synthetic peptide  
US-10-040-128-1

Query Match  
Best Local Similarity 79.3%; Score 73; DB 13; Length 15;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DWINOMHIDEKNT 14  
|||||  
Db 1 DWINOMHIDEKNT 14

RESULT 10  
US-10-243-552-596

; Sequence 596; Application US/10243552  
; Publication No. US20030224379A1  
; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Zhiwei

; APPLICANT: Meng, Gezhi

; TITLE OF INVENTION: Novel Nucleic Acids and

; FILE REFERENCE: 807A  
; CURRENT APPLICATION NUMBER: US/10/243,552

; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 60/322,511

; PRIOR FILING DATE: 2001-09-13  
; PRIOR APPLICATION NUMBER: PCT/US00/35017

; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/488,725

; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 09/552,317

; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: PCT/US01/02623

; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: US 09/491,404

; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: PCT/US01/03800

; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/496,914

; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: US 09/560,875

; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US01/04927

; PRIOR FILING DATE: 2001-02-26  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 998  
; SOFTWARE: pt\_Fl\_genes Version 5.0  
; SEQ ID NO 596

; LENGTH: 167  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-243-552-596

Query Match	62.0%;	Score 57;	DB 12;	Length 167;
Best Local Similarity	57.1%;	Pred. No. 1.1;		
Matches	8;	Conservative	2;	Mismatches 4;
				Indels 0;
				Gaps 0

```
QY      2 W I W N Q M H I D E E K N T      15
          | : | | | : | | | |
Db      45 W W N Q F V P E E M N T      58
```

RESULT 11

US-09-978-295A-284  
Sequence 264, Application US/09978295A  
Patent No. US20020156006A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnovers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Godowsky, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630P1C11  
CURRENT APPLICATION NUMBER: US/09/978,295A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936

1	PRIOR FILING DATE: 1998-03-20
2	PRIOR APPLICATION NUMBER: 60/078910
3	PRIOR FILING DATE: 1998-03-20
4	PRIOR APPLICATION NUMBER: 60/078939
5	PRIOR FILING DATE: 1998-03-20
6	PRIOR APPLICATION NUMBER: 60/079294
7	PRIOR FILING DATE: 1998-03-25
8	PRIOR APPLICATION NUMBER: 60/079656
9	PRIOR FILING DATE: 1998-03-26
10	PRIOR APPLICATION NUMBER: 60/079664
11	PRIOR FILING DATE: 1998-03-27
12	PRIOR APPLICATION NUMBER: 60/079669
13	PRIOR FILING DATE: 1998-03-27
14	PRIOR APPLICATION NUMBER: 60/079663
15	PRIOR FILING DATE: 1998-03-27
16	PRIOR APPLICATION NUMBER: 60/079728
17	PRIOR FILING DATE: 1998-03-27
18	PRIOR APPLICATION NUMBER: 60/079786
19	PRIOR FILING DATE: 1998-03-27
20	PRIOR APPLICATION NUMBER: 60/079920
21	PRIOR FILING DATE: 1998-03-30
22	PRIOR APPLICATION NUMBER: 60/079923
23	PRIOR FILING DATE: 1998-03-30
24	PRIOR APPLICATION NUMBER: 60/080105
25	PRIOR FILING DATE: 1998-03-31
26	PRIOR APPLICATION NUMBER: 60/080107
27	PRIOR FILING DATE: 1998-03-31
28	PRIOR APPLICATION NUMBER: 60/080165
29	PRIOR FILING DATE: 1998-03-31
30	PRIOR APPLICATION NUMBER: 60/080194
31	PRIOR FILING DATE: 1998-03-31
32	PRIOR APPLICATION NUMBER: 60/080327
33	PRIOR FILING DATE: 1998-04-01
34	PRIOR APPLICATION NUMBER: 60/080328
35	PRIOR FILING DATE: 1998-04-01
36	PRIOR APPLICATION NUMBER: 60/080333
37	PRIOR FILING DATE: 1998-04-01
38	PRIOR APPLICATION NUMBER: 60/080334
39	PRIOR FILING DATE: 1998-04-01
40	PRIOR APPLICATION NUMBER: 60/081070
41	PRIOR FILING DATE: 1998-04-08
42	PRIOR APPLICATION NUMBER: 60/081049
43	PRIOR FILING DATE: 1998-04-08
44	PRIOR APPLICATION NUMBER: 60/081071
45	PRIOR FILING DATE: 1998-04-08
46	PRIOR APPLICATION NUMBER: 60/081195
47	PRIOR FILING DATE: 1998-04-08
48	PRIOR APPLICATION NUMBER: 60/081203
49	PRIOR FILING DATE: 1998-04-09
50	PRIOR APPLICATION NUMBER: 60/081229
51	PRIOR FILING DATE: 1998-04-09
52	PRIOR APPLICATION NUMBER: 60/081555
53	PRIOR FILING DATE: 1998-04-15
54	PRIOR APPLICATION NUMBER: 60/081817
55	PRIOR FILING DATE: 1998-04-15
56	PRIOR APPLICATION NUMBER: 60/081819
57	PRIOR FILING DATE: 1998-04-15
58	PRIOR APPLICATION NUMBER: 60/081952
59	PRIOR FILING DATE: 1998-04-15
60	PRIOR APPLICATION NUMBER: 60/081838
61	PRIOR FILING DATE: 1998-04-15
62	PRIOR APPLICATION NUMBER: 60/082568
63	PRIOR FILING DATE: 1998-04-21
64	PRIOR APPLICATION NUMBER: 60/082569
65	PRIOR FILING DATE: 1998-04-21
66	PRIOR APPLICATION NUMBER: 60/082704
67	PRIOR FILING DATE: 1998-04-22
68	PRIOR APPLICATION NUMBER: 60/082804
69	PRIOR FILING DATE: 1998-04-22
70	PRIOR APPLICATION NUMBER: 60/082700
71	PRIOR FILING DATE: 1998-04-22
72	PRIOR APPLICATION NUMBER: 60/082799
73	PRIOR FILING DATE: 1998-04-22

```

1 PRIOR APPLICATION NUMBER: 60/082796
2 PRIOR FILING DATE: 1998-04-23
3 PRIOR APPLICATION NUMBER: 60/083336
4 PRIOR FILING DATE: 1998-04-27
5 PRIOR APPLICATION NUMBER: 60/083322
6 PRIOR FILING DATE: 1998-04-28
7 PRIOR APPLICATION NUMBER: 60/083392
8 PRIOR FILING DATE: 1998-04-29
9 PRIOR APPLICATION NUMBER: 60/083435
10 PRIOR FILING DATE: 1998-04-29
11 PRIOR APPLICATION NUMBER: 60/083496
12 PRIOR FILING DATE: 1998-04-29
13 PRIOR APPLICATION NUMBER: 60/083499
14 PRIOR FILING DATE: 1998-04-29
15 PRIOR APPLICATION NUMBER: 60/083545
16 PRIOR FILING DATE: 1998-04-29
17 PRIOR APPLICATION NUMBER: 60/083554
18 PRIOR FILING DATE: 1998-04-29
19 PRIOR APPLICATION NUMBER: 60/083558
20 PRIOR FILING DATE: 1998-04-29
21 PRIOR APPLICATION NUMBER: 60/083559
22 PRIOR FILING DATE: 1998-04-29
23 PRIOR APPLICATION NUMBER: 60/083500
24 PRIOR FILING DATE: 1998-04-29
25 PRIOR APPLICATION NUMBER: 60/083742
26 PRIOR FILING DATE: 1998-04-30
27 PRIOR APPLICATION NUMBER: 60/084366
28 PRIOR FILING DATE: 1998-05-05
29 PRIOR APPLICATION NUMBER: 60/084414
30 PRIOR FILING DATE: 1998-05-06
31 PRIOR APPLICATION NUMBER: 60/084441
32 PRIOR FILING DATE: 1998-05-06
33 PRIOR APPLICATION NUMBER: 60/084637
34 PRIOR FILING DATE: 1998-05-07
35 PRIOR APPLICATION NUMBER: 60/084639
36 PRIOR FILING DATE: 1998-05-07
37 PRIOR APPLICATION NUMBER: 60/084640
38 PRIOR FILING DATE: 1998-05-07
39 PRIOR APPLICATION NUMBER: 60/084698
40 PRIOR FILING DATE: 1998-05-07
41 PRIOR APPLICATION NUMBER: 60/084600
42 PRIOR FILING DATE: 1998-05-07
43 PRIOR APPLICATION NUMBER: 60/084627
44 PRIOR FILING DATE: 1998-05-07
45 PRIOR APPLICATION NUMBER: 60/084443
46 PRIOR FILING DATE: 1998-05-07
47 PRIOR APPLICATION NUMBER: 60/085339
48 PRIOR FILING DATE: 1998-05-13
49 PRIOR APPLICATION NUMBER: 60/085338
50 PRIOR FILING DATE: 1998-05-13
51 PRIOR APPLICATION NUMBER: 60/085323
52 PRIOR FILING DATE: 1998-05-13
53 PRIOR APPLICATION NUMBER: 60/085582
54 PRIOR FILING DATE: 1998-05-15
55 PRIOR APPLICATION NUMBER: 60/085700
56 PRIOR FILING DATE: 1998-05-15
57 PRIOR APPLICATION NUMBER: 60/085689
58 PRIOR FILING DATE: 1998-05-15
59 PRIOR APPLICATION NUMBER: 60/085579
60 PRIOR FILING DATE: 1998-05-15
61 PRIOR APPLICATION NUMBER: 60/085580
62 PRIOR FILING DATE: 1998-05-15
63 PRIOR APPLICATION NUMBER: 60/085573
64 PRIOR FILING DATE: 1998-05-15
65 PRIOR APPLICATION NUMBER: 60/085704
66 PRIOR FILING DATE: 1998-05-15
67 PRIOR APPLICATION NUMBER: 60/085697

```

Query March	62.0%;	Score 57;	DB 9;	Length 772;
Best Local Similarity	57.1%;	Pred. No. 4.6;		
Matches	8;	Conservative	2;	Mismatches 4;
				Indels 0;
				Gaps 0;
QY	2	WIMWQHAIDEEKNT	15	

```

Db          45 WYMNQFVPEEMT 58      ||| : |||
RESULT 12
US-09-978-697-264
/ Sequence 264, Application US/0978697
/ Patent No. US20020169264A1
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi
/ APPLICANT: Baker Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnovers, Luc
/ APPLICANT: Eaton, Dan
/ APPLICANT: Ferrara, Napoleon
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Kljavin, Ivar J.
/ APPLICANT: Kuo, Sophia S.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James;
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Shelton, David L.
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ FILE REFERENCE: P2630P1C27
/ CURRENT APPLICATION NUMBER: US/09/978,697
/ CURRENT FILING DATE: 2001-10-16
/ PRIOR APPLICATION NUMBER: 09/918585
/ PRIOR FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/064249
/ PRIOR FILING DATE: 1997-11-03
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066364
/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/077450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077641
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077791
/ PRIOR FILING DATE: 1998-03-12
/ PRIOR APPLICATION NUMBER: 60/078004
/ PRIOR FILING DATE: 1998-03-13
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078936
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078910
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079294
/ PRIOR FILING DATE: 1998-03-25

```

PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392

PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 62.0%; Score 57; DB 9; Length 772;  
Best Local Similarity 57.1%; Pred. No. 4.6;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 WYNNQWHDKEKNT 15  
Db 45 WYNNQWHDKEKNT 58

RESULT 13  
US-09-978-192A-264  
; Sequence 264, Application US/09978192A

Patent No. US2002017553A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnovers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Kujavian, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
ACIDS  
FILE REFERENCE: P2630P1C9  
CURRENT APPLICATION NUMBER: US/09/978,192A  
PRIOR FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/07450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078866  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663

PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29

```

1 PRIOR APPLICATION NUMBER: 60/083545
2 PRIOR FILING DATE: 1998-04-29
3 PRIOR APPLICATION NUMBER: 60/083554
4 PRIOR FILING DATE: 1998-04-29
5 PRIOR APPLICATION NUMBER: 60/083558
6 PRIOR FILING DATE: 1998-04-29
7 PRIOR APPLICATION NUMBER: 60/083559
8 PRIOR FILING DATE: 1998-04-29
9 PRIOR APPLICATION NUMBER: 60/083500
10 PRIOR FILING DATE: 1998-04-29
11 PRIOR APPLICATION NUMBER: 60/083742
12 PRIOR FILING DATE: 1998-04-30
13 PRIOR APPLICATION NUMBER: 60/084366
14 PRIOR FILING DATE: 1998-05-05
15 PRIOR APPLICATION NUMBER: 60/084414
16 PRIOR FILING DATE: 1998-05-06
17 PRIOR APPLICATION NUMBER: 60/084441
18 PRIOR FILING DATE: 1998-05-06
19 PRIOR APPLICATION NUMBER: 60/084637
20 PRIOR FILING DATE: 1998-05-07
21 PRIOR APPLICATION NUMBER: 60/084639
22 PRIOR FILING DATE: 1998-05-07
23 PRIOR APPLICATION NUMBER: 60/084640
24 PRIOR FILING DATE: 1998-05-07
25 PRIOR APPLICATION NUMBER: 60/084598
26 PRIOR FILING DATE: 1998-05-07
27 PRIOR APPLICATION NUMBER: 60/084600
28 PRIOR FILING DATE: 1998-05-07
29 PRIOR APPLICATION NUMBER: 60/084627
30 PRIOR FILING DATE: 1998-05-07
31 PRIOR APPLICATION NUMBER: 60/084643
32 PRIOR FILING DATE: 1998-05-07
33 PRIOR APPLICATION NUMBER: 60/085339
34 PRIOR FILING DATE: 1998-05-13
35 PRIOR APPLICATION NUMBER: 60/085338
36 PRIOR FILING DATE: 1998-05-13
37 PRIOR APPLICATION NUMBER: 60/085323
38 PRIOR FILING DATE: 1998-05-13
39 PRIOR APPLICATION NUMBER: 60/085582
40 PRIOR FILING DATE: 1998-05-15
41 PRIOR APPLICATION NUMBER: 60/085700
42 PRIOR FILING DATE: 1998-05-15
43 PRIOR APPLICATION NUMBER: 60/085689
44 PRIOR FILING DATE: 1998-05-15
45 PRIOR APPLICATION NUMBER: 60/085579
46 PRIOR FILING DATE: 1998-05-15
47 PRIOR APPLICATION NUMBER: 60/085580
48 PRIOR FILING DATE: 1998-05-15
49 PRIOR APPLICATION NUMBER: 60/085573
50 PRIOR FILING DATE: 1998-05-15
51 PRIOR APPLICATION NUMBER: 60/085704
52 PRIOR FILING DATE: 1998-05-15
53 PRIOR APPLICATION NUMBER: 60/085697
54

```

```

Query Match      62.0%; Score 57; DB 9; Length 772;
Best Local Similarity 57.1%; Pred. No. 4.6;
Matches      8; Conservative      2; Mismatches      4; Indels      0; Gaps      0;

QY      2 WIMQNMHIDEKNT 15
      |:| | | : | | |
Db      45 WWMNCFVPEEMNT 58

RESULT 14
US-09-999-832A-264
; Sequence 264, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Raton, Dan

```

APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Guiney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijaviri, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
ACIDS OR INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630P1C63  
CURRENT APPLICATION NUMBER: US/09/999,832A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 09/948585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079556  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30

```

; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match      62.0%; Score 57; DB 9; Length 772;
Best Local Similarity 57.1%; Pred. No. 4.6;
Matches      8; Conservative      2; Mismatches      4; Indels      0; Gaps      0;

Qy      2 WWMQMHIDEKNT 15
Db      45 WWMQFVPEEMNT 58

RESULT 15
; Sequence 264; Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gottlieb, Mary E.
; APPLICANT: Goddard, Audrey

```

APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, V. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar T.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumaw, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C7  
CURRENT APPLICATION NUMBER: US/09/978,189  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/074450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/07791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165

PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05



```

; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match      62.0%; Score 57; DB 10; Length 772;
Best Local Similarity 57.1%; Pred. No. 4.6;
Matches      8; Conservative      2; Mismatches      4; Indels      0; Gaps      0;

Qy      2 WTWNQMHIDEKXT 15
      ||||| : |||
Db      45 WTWNQFVPEEMT 58

```

Search completed: July 29, 2004, 11:50:33  
 Job time : 39.3333 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:30:30 ; Search time 6.66667 Seconds

(without alignments)

117.158 Million cell updates/sec

Title: US-10-040-128-2

Sequence: 92  
1 DWIMQMHIDEKNE 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	784	1	CAD5_MOUSE
2	87	94.6	782	1	CAD5_PIG
3	87	94.6	784	1	CAD5_HUMAN
4	52	56.5	772	1	CAD7_HUMAN
5	49	53.3	799	1	CAD8_HUMAN
6	49	53.3	799	1	CAD8_MOUSE
7	49	53.3	799	1	CAD8_RAT
8	47.5	51.6	388	1	YDRA_BACSU
9	45	48.9	792	1	CAD8_CHICK
10	45	48.9	796	1	CAD8_MOUSE
11	45	48.9	796	1	CAD8_MOUSE
12	45	48.9	813	1	CAD7_MOUSE
13	45	48.9	813	1	CAD7_MOUSE
14	45	48.9	813	1	CAD7_MOUSE
15	45	48.9	813	1	CAD7_MOUSE
16	44	47.8	785	1	CAD7_CHICK
17	44	47.8	785	1	CAD7_CHICK
18	44	47.8	790	1	CAD7_HUMAN
19	44	47.8	794	1	CAD7_HUMAN
20	44	47.8	801	1	CAD7_HUMAN
21	43	46.7	435	1	HISX_BUCAI
22	42	45.7	427	1	CRSY_SALTY
23	41.5	45.1	789	1	NIXE_ECOLI
24	41	44.6	314	1	NIXE_ECOLI
25	41	44.6	363	1	DR3X_BUCAI
26	41	44.6	547	1	SPAK_HUMAN
27	41	44.6	553	1	SPAK_HUMAN
28	41	44.6	556	1	SPAK_HUMAN
29	41	44.6	788	1	CADA_HUMAN
30	41	44.6	789	1	CAD6_RAT
31	41	44.6	789	1	CAD6_RAT
32	41	44.6	790	1	CAD6_CHICK
33	41	44.6	790	1	CAD6_HUMAN

34	40.5	44.0	230	1	T2E7_ECOLI	P50194 escherichia
35	40.5	44.0	757	1	IKKB_MOUSE	O98351 mus musculus
36	40.5	44.0	757	1	IKKB_RAT	O99718 ratius norv
37	40	43.5	146	1	HBBC_HOPI	P82316 hoplosternu
38	40	43.5	373	1	LYSI_YEAST	P38998 saccharomyc
39	40	43.5	418	1	GLYA_STRP3	O8K798 streptococc
40	40	43.5	418	1	GLYA_STRP8	O8P122 streptococc
41	40	43.5	433	1	GLYA_STRP5	O96221 streptococc
42	40	43.5	433	1	UROC_MOUSE	P06869 mus musculus
43	40	43.5	475	1	PBM2_ARATH	O94410 arabidopsi
44	40	43.5	508	1	TNSD_ECOLI	P13991 escherichia
45	40	43.5	560	1	CALX_SCHPO	P35581 schistosach

## ALIGNMENTS

RESULT 1  
CAD5\_MOUSE STANDARD; PRT; 784 AA.  
ID CAD5\_MOUSE  
AC P55284; O35542;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_Taxid=10090;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain capillary;  
RX MEDLINE=96141083; PubMed=8555485;  
RA Breier G., Breviario F., Cavada L., Berthier R., Schnerch H.,  
RT Gotsch U., Westweber D., Risau W., Dejana E.,  
RT "Molecular cloning and expression of murine vascular endothelial-  
RT cadherin in early stage development of cardiovascular system.";  
RN Blood 87:630-641(1996).  
[2]  
RP REVISIONS TO 67-70.  
RA Breviario F.,  
RN Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A., AND FUNCTION.  
RC STRAIN=BALB/c; TISSUE=Breast carcinoma;  
RX MEDLINE=97364256; PubMed=9220534;  
RA Matsuyoshi N., Toda K.-I., Horiguchi Y., Tanaka T., Nakagawa S.,  
RT Takeichi M., Imanura S.,  
RT "In vivo evidence of the critical role of cadherin-5 in murine  
RT vascular integrity";  
RN Proc. Assoc. Am. Physicians 109:362-371(1997).  
[4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klusner R.D., Collins F.S., Wagner L., Shenmen C.F., Bhat N.K.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stempleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiki S., Carninci P., Prange C.,  
RA Raha S.S., Iqbalano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bostak S.A., McKernan K.J., Garcia A.M., Gay L.J., Huliy S.W.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smialowski D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X83930; CA58782.2; -  
 CC EMBL; D63942; BAA22617.1; -  
 CC EMBL; BC054790; AA554790.1; -  
 CC HSSP; P15116; INCI  
 CC MGD; MGI:105057; Cdh5.  
 CC InterPro; IPR002126; Cadherin.  
 CC InterPro; IPR000233; Cadherin\_C\_term.  
 CC Pfam; PF00028; cadherin; 5.  
 CC Pfam; PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS; PR00205; CADHERIN.  
 CC SMART; SM00112; CA; 5.  
 CC PROSITE; PS00232; CADHERIN\_1; 3.  
 CC PROSITE; PS0268; CADHERIN\_2; 5.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT PROPEP 25 45 POTENTIAL.  
 FT CHAIN 46 784 VASCULAR ENDOTHELIAL-CADHERIN.  
 FT DOMAIN 46 593 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 594 620 POTENTIAL.  
 FT DOMAIN 621 784 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 46 149 CADHERIN 1.  
 FT DOMAIN 150 256 CADHERIN 2.  
 FT DOMAIN 257 371 CADHERIN 3.  
 FT DOMAIN 372 476 CADHERIN 4.  
 FT DOMAIN 477 593 CADHERIN 5.  
 FT DOMAIN 738 753 SER-RICH.  
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 784 AA; 87902 MW; 7B75698DE2F7E160 CRC64;  
 QY Query Match 100.0%; Score 92; DB 1; Length 784;  
 QY Best Local Similarity 100.0%; Pred. No. 8.5e-07;  
 Db Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 2  
 ID CAD5\_PIG STANDARD; PRT; 782 AA.  
 AC 002840;  
 DT 15-UTR-1998 (Rel. 36, Created)  
 DT 15-UTR-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5).  
 GN Cdh5.  
 OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sui.  
 OC NCBI\_Taxid=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kishaw P.J.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions. It  
 CC associates with alpha-catenin forming a link to the cytoskeleton  
 CC (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Found at cell-cell  
 CC boundaries and probably at cell-matrix boundaries (By similarity).  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Y13919; CAA74225.1; -  
 CC HSSP; P09603; IBDH.  
 CC InterPro; IPR002126; Cadherin.  
 CC InterPro; IPR000233; Cadherin\_C\_term.  
 CC Pfam; PF00028; cadherin; 5.  
 CC Pfam; PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS; PR00205; CADHERIN.  
 CC SMART; SM00112; CA; 5.  
 CC PROSITE; PS00232; CADHERIN\_1; 3.  
 CC PROSITE; PS0268; CADHERIN\_2; 5.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT PROPEP 23 44 POTENTIAL.  
 FT CHAIN 45 782 VASCULAR ENDOTHELIAL-CADHERIN.  
 FT DOMAIN 45 592 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 593 619 POTENTIAL.  
 FT DOMAIN 620 782 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 45 148 CADHERIN 1.  
 FT DOMAIN 149 255 CADHERIN 2.  
 FT DOMAIN 256 370 CADHERIN 3.  
 FT DOMAIN 371 475 CADHERIN 4.  
 FT DOMAIN 476 592 CADHERIN 5.  
 FT DOMAIN 736 751 SER-RICH.  
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 782 AA; 87546 MW; 7403F974E2DF782F CRC64;  
 QY Query Match 94.6%; Score 87; DB 1; Length 782;  
 QY Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
 Db Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 3  
 ID CAD5\_HUMAN STANDARD; PRT; 784 AA.  
 AC P33151;  
 DT 01-OCT-1993 (Rel. 27, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5)  
 DE (7B4 antigen) (CD144 antigen).  
 GN CDH5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Endothelial cells;  
 RX MEDLINE=95353875; PubMed=7627717;  
 RA Breviaro F., Cavada L., Corada M., Martin-Padura I., Navarro P.,  
 RA Gelay J., Introna M., Gulino D., Lampugnani M.G., Dejana E.;  
 RT "Functional properties of human vascular endothelial cadherin  
 (7B4/cadherin-5), an endothelium-specific cadherin.";  
 RL Atheroscler. Thromb. Vasc. Biol. 15:1229-1239(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=97362755; PubMed=9219219;  
 RA Ali J., Liao F., Martens E., Muller W.A.;  
 RT "Vascular endothelial cadherin (VE-cadherin): cloning and role in  
 RL Microcirculation cell-cell adhesion.";  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC PubMed=10861224;  
 RA Shimoyama Y., Tsujimoto G., Kitejima M., Natori M.;  
 RT "Identification of three human type-II classic cadherins and frequent  
 RT heterophilic interactions between different subclasses of type-II  
 RT classic cadherins.";  
 RL Biochem. J. 349:159-167(2000).  
 RN [4]  
 RP SEQUENCE OF 5-784 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=91283540; PubMed=2059658;  
 RA Suzuki S., Sano K., Tanihara H.;  
 RT "Diversity of the cadherin family: evidence for eight new cadherins  
 RT in nervous tissue.";  
 RL Cell Regul. 2:261-270(1991).  
 RN [5]  
 RP PARTIAL SEQUENCE.  
 RC TISSUE=Endothelial cells;  
 RX MEDLINE=92394977; PubMed=1522121;  
 RA Lampugnani M.G., Resnati M., Rafteri M., Pigott R., Pisacane A.,  
 RA Hoven G., Ruco L.P., Dejana E.;  
 RT "A novel endothelial-specific membrane protein is a marker of  
 RT cell-cell contacts.";  
 RL J. Cell Biol. 118:1511-1522(1992).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions. It  
 CC associates with alpha-catenin forming a link to the cytoskeleton.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Found at cell-cell  
 CC boundaries and probably at cell-matrix boundaries.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -1- DATABASE: NAME=PROV, NOTE=CD guide CD144 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd144.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; X79981; CA56306.1; -.  
 DR EMBL; U84722; AAB41796.1; -.  
 DR EMBL; AB035304; BAA87418.1; -.  
 DR EMBL; X59796; CAA42468.1; -.  
 DR PIR; S49893; ITHUC5.  
 DR HSSP; P15116; INCU.  
 DR GlycoSuiteDB; P31511; -.  
 DR Genew; HGNC:1764; CDH5.  
 DR MIM; 601120; -.  
 DR GO; GO:0005624; C:membrane fraction; TAS.  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0007156; P:homophilic cell adhesion; TAS.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR000233; Cadherin\_C\_term.  
 DR Pfam; PF00028; cadherin; 5.  
 DR Pfam; PF01049; Cadherin\_C\_term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS50268; CADHERIN\_2; 5.  
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KM Signal.  
 FT SIGNAL. 1 25  
 FT PROPEP 26 47  
 FT CHAIN 48 784  
 FT DOMAIN 48 593  
 FT TRANSMEM 594 620  
 FT DOMAIN 621 784  
 FT DOMAIN 48 151  
 FT DOMAIN 152 258  
 FT DOMAIN 259 372  
 FT DOMAIN 373 477  
 FT DOMAIN 478 593  
 FT DOMAIN 593 736  
 FT CARBOHYD 61 61  
 FT CARBOHYD 112 112  
 FT CARBOHYD 157 157  
 FT CARBOHYD 362 362  
 FT CARBOHYD 442 442  
 FT CARBOHYD 523 523  
 FT CARBOHYD 535 535  
 FT CONFLICT 517 517  
 SQ SEQUENCE 784 AA; 87516 MW; F643BFC22A599DE CRC64;  
 Query Match 94.6%; Score 87; DB 1; Length 784;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DWIMQMHIDEKN 14  
 Db 48 DWIMQMHIDEKN 61  
 ID CADU HUMAN STANDARD; PRT; 772 AA.  
 AC Q9H159;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Cadherin-19 precursor (UNQ478/PRO941).  
 GN CDH19 OR CDH7L2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20453188; PubMed=10995570;  
 RX Kools P., Van Imschoot G., van Roy F.;  
 RT "Characterization of three novel human cadherin genes (CDH7, CDH19,  
 RT and CDH20) clustered on chromosome 18q22-q23 and with high homology  
 RT to chicken cadherin-7.";

RL Genomics 68:283-295(2000).

CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed in many tissues, with the exception  
 CC of uterus.

CC -1- SIMILARITY: Contains 5 cadherin domains.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AJ007607; CAC13126.1; -  
 CC EMBL: AY358654; AA089017.1; -  
 CC HSSP: P15116; INCU.  
 CC Genew: HGNC:1758; CDH19.  
 CC MIM: 603016; -  
 CC GO: GO:0016021; C: integral to membrane; NAS.  
 CC GO: GO:0007156; P: homophilic cell adhesion; NAS.  
 CC InterPro: IPR002126; Cadherin.  
 CC InterPro: IPR00233; Cadherin\_C\_term.  
 CC Pfam: PF00028; cadherin\_5.  
 CC Pfam: PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS: PR00205; CADHERIN.  
 CC SMART: SM0112; CA; 5.  
 CC PROSITE: PS00232; CADHERIN\_1; 2.  
 CC PROSITE: PS0268; CADHERIN\_2; 5.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.

FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 43 POTENTIAL.  
 FT CHAIN 44 772 CADHERIN-19.  
 FT DOMAIN 44 596 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 597 617 POTENTIAL.  
 FT DOMAIN 618 772 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 149 256 CADHERIN 1.  
 FT DOMAIN 257 370 CADHERIN 2.  
 FT DOMAIN 371 470 CADHERIN 3.  
 FT DOMAIN 470 581 CADHERIN 4.  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 508 508 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 516 516 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 772 AA; 87002 MW; 650AD27480343C39 CRC64;

Query Match 56.5%; Score 52; DB 1; Length 772;  
 Best Local Similarity 53.8%; Pred. No. 1.6;  
 Matches 7; Conservativity 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 WIMNOMHIDEKN 14  
 DB 45 WWMNQFVPEKN 57

RESULT 5  
 CAD8\_HUMAN STANDARD; PRT; 799 AA.  
 AC P55286; Q9ULB2;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cadherin-8 precursor.

GN CDH8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RX PubMed=10861224;  
 RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;  
 RT "Identification of three human type-II classic cadherins and frequent  
 RT heterophilic interactions between different subclasses of type-II  
 RT classic cadherins.";  
 RL Biochem. J. 349:159-167(2000).  
 RN [2]  
 RP SEQUENCE OF 7-799 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95073006; PubMed=7982033;  
 RA Tanihara H., Sano K., Helmark R.L., St John T., Suzuki S.;  
 RT "Cloning of five human cadherins clarifies characteristic features of  
 RT cadherin extracellular domain and provides further evidence for two  
 RT structurally different types of cadherin.";  
 RL Cell Adhes. Commun. 2:15-26(1994).  
 RN [3]  
 RP SEQUENCE OF 294-799 FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=91283540; PubMed=2059658;  
 RA Suzuki S., Sano K., Tanihara H.;  
 RT "Diversity of the cadherin family: evidence for eight new cadherins  
 RT in nervous tissue.";  
 RL Cell Regul. 2:261-270(1991).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Mainly expressed in brain. Found in certain  
 CC nerve cell lines, such as retinoblasts, glioma cells and  
 CC neuroblasts.

CC -1- SIMILARITY: Contains 5 cadherin domains.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AB035305; BA87419.1; -  
 CC EMBL: U34060; AA035628.1; ADT\_INIT.  
 CC HSSP: P15116; INCU.  
 CC Genew: HGNC:1767; CDH8.  
 CC MIM: 603008; -  
 CC GO: GO:007155; P: cell adhesion; TAS.  
 CC InterPro: IPR002126; Cadherin.  
 CC InterPro: IPR00233; Cadherin\_C\_term.  
 CC Pfam: PF00028; cadherin\_5.  
 CC Pfam: PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS: PR00205; CADHERIN.  
 CC SMART: SM00112; CA; 5.  
 CC PROSITE: PS00232; CADHERIN\_1; 3.  
 CC PROSITE: PS0268; CADHERIN\_2; 5.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.

FT SIGNAL 1 29 POTENTIAL.  
 FT PROPEP 30 61 POTENTIAL.  
 FT CHAIN 62 799 CADHERIN-8.  
 FT DOMAIN 62 621 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 622 642 POTENTIAL.  
 FT DOMAIN 643 799 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 62 167 CADHERIN 1.  
 FT DOMAIN 168 276 CADHERIN 2.  
 FT DOMAIN 277 391 CADHERIN 3.

```

FT DOMAIN 392 494 CADHERIN 4.
FT DOMAIN 495 616 CADHERIN 5.
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 647 647 V -> D (IN REF. 2 AND 3).
FT CONFLICT 799 AA; 88253 MM; 9B119B86039C6A0A CRC64;
SQ SEQUENCE

```

```

Query Match 53.3%; Score 49; DB 1; Length 799;
Best Local Similarity 63.6%; Pred. No. 4.9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 WWMNQMHIDE 12
Db 63 WWMNQMFVLEE 73

```

## RESULT 6

```

CAD8_MOUSE STANDARD; PRT; 799 AA.
ID CAD8_MOUSE
P97251;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-8 precursor.
GN CDH8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster / NIH;
RX MEDLINE=97174321; PubMed=9022055;
RA Korematsu K., Redies C.;
RT "Restricted expression of cadherin-8 in segmental and functional
subdivisions of the embryonic mouse brain.";
RL Dev. Dyn. 208:178-189(1997).

```

```

CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 5 cadherin domains.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; X95600; CAA64857.1; -
DR HSSP; P15116; INCU.
DR MCD; MGI:107434; Cdh8.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SMO0112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS50268; CADHERIN_2; 5.
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL. 1 29 POTENTIAL.
FT PROPEP 30 61 POTENTIAL.
FT CHAIN 62 799 CADHERIN-8.

```

```

FT DOMAIN 62 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 642 POTENTIAL.
FT DOMAIN 643 643 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 62 167 CADHERIN 1.
FT DOMAIN 168 276 CADHERIN 2.
FT DOMAIN 277 391 CADHERIN 3.
FT DOMAIN 392 494 CADHERIN 4.
FT DOMAIN 495 616 CADHERIN 5.
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE

```

```

Query Match 53.3%; Score 49; DB 1; Length 799;
Best Local Similarity 63.6%; Pred. No. 4.9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 WWMNQMHIDE 12
Db 63 WWMNQMFVLEE 73

```

## RESULT 7

```

CAD8_RAT STANDARD; PRT; 799 AA.
ID CAD8_RAT
O54801;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cadherin-8 precursor.
GN CDH8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=98190518; PubMed=9521872;
RA Kido M., Obata S., Tanihara H., Koehelle J.M., Seldin M.F.,
RA Taketani S., Suzuki S.T.;
RT "Molecular properties and chromosomal location of cadherin-8.";
RL Genomics 48:186-194(1998).

```

```

CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O54800-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O54800-2; Sequence=VSP_000638, VSP_000639;

```

```

CC -!- SIMILARITY: Contains 5 cadherin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AB010436; BAA24452.1; -
DR EMBL; AB010437; BAA24453.1; -
DR HSSP; P15116; INCU.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.

```

DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN 1; 3.  
 DR PROSITE; PS00268; CADHERIN 2; 5.  
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 Signal; Alternative splicing  
 FT SIGNAL 1 29 POTENTIAL.  
 FT PROPEP 30 61 POTENTIAL.  
 FT CHAIN 62 799 CADHERIN-8.  
 FT DOMAIN 62 621 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 622 642 POTENTIAL.  
 FT DOMAIN 643 799 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 62 167 CADHERIN 1.  
 FT DOMAIN 168 276 CADHERIN 2.  
 FT DOMAIN 277 391 CADHERIN 3.  
 FT DOMAIN 392 494 CADHERIN 4.  
 FT DOMAIN 495 616 CADHERIN 5.  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 514 532 IOTVSADKDDPKNGHFPL -> NISMLILNMEVYNCFLV  
 FT VARSPLIC 514 532 N (in isoform 2).  
 FT VARSPLIC 533 799 /FtId=VSP 000638.  
 FT VARSPLIC 533 799 /FtId=VSP 000639.  
 SQ SEQUENCE 799 AA; 88332 MW; F01D145F80966CB6 CRC64;  
 Query Match 53.3%; Score 49; DB 1; Length 799;  
 Best Local Similarity 63.6%; Pred. No. 4.9;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 WINNOMHI-DEEKN 12  
 Db 63 WINNOMHI-DEEKN 73  
 RESULT 8  
 YUBA\_BACSU STANDARD; PRT; 388 AA.  
 AC 032086;  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9394377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 Bortiss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 Brouillet S., Bruschi C.V., Caldwell B., Capiano V., Carter N.M.,  
 Choi S.K., Codani J.U., Comerton I.F., Cummings N.J., Daniel R.A.,  
 Denzot F., Devilling K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 Frits C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,  
 Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 Hillbert H., Holsappel S., Hosono S., Hullu M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 Kobayashi Y., Koester P., Konigstein G., Krogh S., Kumano M.,  
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
 Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,  
 Paro V., Pohl T.M., Portecelle B., Portwolk S., Prescott A.M.,  
 Pessacq E., Pujic P., Purnelle D., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,  
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Taccanti E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tognoni A.,  
 RA Toso V., Uchiyama S., Vandenbol M., Vanier F., Vassart A.,  
 RA Vairi A., Wambut R., Wedler E., Wedler H., Weikensger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis";  
 RT Nature 390:249-256(1997).  
 RL SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE UPF018 (PERM) FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE UPF018 (PERM) FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; Z99119; CAB15094.1; -.  
 DR PIR; D70006; D70006.  
 DR Subtilisin; B61350; ynda.  
 DR InterPro; IPR002549; UPF0118.  
 DR Pfam; PF01594; UPF0118; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 20 40 POTENTIAL.  
 FT TRANSMEM 47 67 POTENTIAL.  
 FT TRANSMEM 86 106 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 FT TRANSMEM 237 257 POTENTIAL.  
 FT TRANSMEM 262 282 POTENTIAL.  
 FT TRANSMEM 286 306 POTENTIAL.  
 FT TRANSMEM 331 351 POTENTIAL.  
 SQ SEQUENCE 388 AA; 43610 MW; EFAB74A052C582AE CRC64;  
 Query Match 51.6%; Score 47.5; DB 1; Length 388;  
 Best Local Similarity 62.5%; Pred. No. 4;  
 Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;  
 QY 1 WINNOMHI-DEEKN 15  
 Db 366 DFKXQSHYDEKNE 381  
 RESULT 9  
 CADB\_CHICK STANDARD; PRT; 792 AA.  
 ID CADB\_CHICK  
 AC 093319;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last annotation update)  
 DE Cadherin-11 precursor.  
 GN Gallus gallus (Chicken).  
 OS Gallus gallus (Chicken).  
 OC Chordata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Chordata; Galliformes; Phasianidae; Phasianinae;  
 OC Archaeosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White Leghorn;  
 RA Wei J., Dong X.R., Topouzis S., Zimmer W.E., Broders F., Thierry J.P.,  
 RA Kotliarsky V., Majesky M.W.;  
 RT "Molecular cloning of chick cadherin 11 and its expression during  
 RT smooth muscle differentiation and formation of the tunica media";  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).





```

FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 384 486 CADHERIN 4.
FT DOMAIN 487 612 CADHERIN 5.
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 632 693 N-FLYFTRARCKEPLIVEEDVEDVITYDEGAGED
TEADRTATONPDGNGFIFR -> GCPSPERPSPREDMR
LTYLFGTQMLFSPYKXNRRCGLGVTKPLFYVATSEPT
LTYSL (in isoform 2).
FT FT Missing (in isoform 2).
FT VARSPLIC 694 796 /FTRID=VSP_000640.
FT CONFLICT 271 272 /FTRID=VSP_000641.
FT CONFLICT 273 273 RL -> SV TIN REF. 2).
FT CONFLICT 340 340 M -> I (IN REF. 2).
FT CONFLICT 373 373 E -> K (IN REF. 2).
FT CONFLICT 471 471 S -> A (IN REF. 2).
FT CONFLICT 471 471 Q -> K (IN REF. 2).
SQ SEQUENCE 796 AA; 88049 MW; 2C67044C78ADB82E CRC64;

Query Match 48.9%; Score 45; DB 1; Length 796;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WINNOMHIDE 12
DB 55 WWMNOFFVIBE 65

RESULT 11
CAB_MOUSE STANDARD; PRT; 796 AA.
AC P55286;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cadherin-11 precursor (Osteoblast-cadherin) (OB-cadherin) (OSF-4).
OS CCH11 OR CAD-11.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=95269886; PubMed=7750649;
RA Hofmann I.H., Balling R.;
RT "Cloning and expression analysis of a novel mesodermally expressed
cadherin.";
RL Dev. Biol. 169:337-346(1995).
RN 2;
RP SEQUENCE FROM N.A.
RX MEDLINE=95269887; PubMed=7750650;
RA Kimura Y., Matsunami H., Inoue T., Shimamura K., Uchida N., Ueno T.,
RA Miyazaki T., Takeichi M.;
RT "Cadherin-11 expressed in association with mesenchymal morphogenesis
in the head, somite, and limb bud of early mouse embryos.";
RL Dev. Biol. 169:347-358(1995).
RN 3;
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Calvaria;
RX MEDLINE=94216322; PubMed=8163513;
RA Okazaki M., Takeshita S., Kawai S., Kikuno R., Tsujimura A., Kudo A.,
RA Amann E.;
RT "Molecular cloning and characterization of OB-cadherin, a new member
of cadherin family expressed in osteoblasts.";
RL J. Biol. Chem. 269:12092-12098(1994).
RN 4;
RP SEQUENCE FROM N.A.
RX TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Fellingsold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallya S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Maller J.A., Guaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smilins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 5;
RP DEVELOPMENTAL STAGE.
RX STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=97033837; PubMed=879495;
RA Munro S.B., Blaschuk O.W.;
RT "A comprehensive survey of the cadherins expressed in the testes of
fetal, immature, and adult mice utilizing the polymerase chain
reaction.";
RL Biol. Reprod. 55:822-827(1996).
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
They preferentially interact with themselves in a homophilic
manner in connecting cells; cadherins may thus contribute to the
sorting of heterogeneous cell types.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Selectively expressed in osteoblastic cell
lines, precursor cell lines of osteoblasts, and primary
osteoblastic cells from calvaria, as well as in lung, testis, and
brain tissues at low levels.
CC -1- DEVELOPMENTAL STAGE: In the testis, expression is highest in fetal
gonad and decreases 8-fold in newborn.
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X77557; CA54674.1; -
DR EMBL, D31963; BAA06730.1; -
DR EMBL, D21253; BAA04797.1; -
DR EMBL, BC046314; AAH46314.1; -
DR PIR, A53584; A53584.
DR PIR, I48277; I48277.
DR PIR, I49556; I49556.
DR HSSP, P15116; INCU.
DR MGD, MGI:99217; Cdh11.
DR GO, GO:0005737; C:cytoplasm; IDA.
DR GO, GO:0005886; C:plasma membrane; IDA.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000233; Cadherin_C_term.
DR Pfam: PF00028; cadherin; 5
DR Pfam: PF01049; Cadherin_C_term; 1.
DR PRINTS, PR00205; CADHERIN.
DR SMART, SM00112; CA; 5.
DR PROSITE, PS00232; CADHERIN 1; 3.
DR PROSITE, PS0268; CADHERIN 2; 5.
KM Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 53 POTENTIAL.
FT CHAIN 54 796 CADHERIN-11.
FT DOMAIN 54 617 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 618 640 POTENTIAL.
FT DOMAIN 641 796 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 54 159 CADHERIN 1.
FT DOMAIN 160 268 CADHERIN 2.
FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 384 486 CADHERIN 4.
FT DOMAIN 487 612 CADHERIN 5.
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 462 462 E -> D (IN REF. 1).
FT CONFLICT 589 589 T -> L (IN REF. 2).
FT CONFLICT 655 655 D -> N (IN REF. 2).
FT CONFLICT 751 751 V -> M (IN REF. 1).
FT CONFLICT 777 777 P -> Q (IN REF. 2).
FT CONFLICT 782 782 L -> P (IN REF. 2).
SQ SEQUENCE 796 AA; 88112 MW; 0D584D24641DD529 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 796;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 2 WIMNQMIDEE 12
Db 55 WWMNQFVVEE 65

RESULT 12
CADM_MOUSE STANDARD; PRT; 813 AA.
ID CADM_MOUSE
AC Q9WTF5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cadherin-22 precursor (PB-cadherin).
GN CDH22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=99326347; PubMed=10398531;
RA Kitajima K., Koshimizu U., Nakamura T.;
RT "Expression of a novel type of classic cadherin, PB-cadherin in
RT developing brain and limb buds."
RT Dev. Dyn. 215:206-214(1999).
RN [1]
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. PB-cadherins may have a role
CC in the morphological organization of pituitary gland and brain
CC tissues.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Predominantly expressed in brain. Abundant in
CC olfactory bulb, cerebrum, and cerebellum, less in pons, medulla,
CC and spinal cord. Low expression in heart. No expression in lung,
CC liver, spleen, kidney, testis, stomach, intestine, colon, and
CC placenta.
CC -1- DEVELOPMENTAL STAGE: Expressed at 9.5 dpc onwards. At 10.5 dpc, in
CC brain (telencephalic vesicles and isthmus), spinal cord and limb
CC buds (in the zone of polarizing activity). At 14.5 dpc, in
CC olfactory bulb and cerebellum.
CC -1- INDUCTION: Down-regulated by thyroid hormone.
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and, for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB019618; BAA34426.1; -

```

```

DR MGD; MGI:1341843; Cdh22.
DR HSSP; P15116; INCU.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN 1; 2.
DR PROSITE; PS00268; CADHERIN 2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 813 CADHERIN-22.
FT DOMAIN 33 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 642 POTENTIAL.
FT DOMAIN 643 813 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 61 165 CADHERIN 1.
FT DOMAIN 166 274 CADHERIN 2.
FT DOMAIN 275 391 CADHERIN 3.
FT DOMAIN 392 495 CADHERIN 4.
FT DOMAIN 496 613 CADHERIN 5.
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 813 AA; 88021 MW; 5510F9848D976567 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 813;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 2 WIMNQMIDEE 12
Db 61 WWMNQFVVEE 71

RESULT 13
CADM_RAT STANDARD; PRT; 813 AA.
ID CADM_RAT
AC Q63315; Q63561;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cadherin-22 precursor (PB-cadherin).
GN CDH22.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RC STRAIN=Wistar; TISSUE=Brain; and Pituitary;
RX MEDLINE=9621232; PubMed=8626716;
RA Sugimoto K., Honda S., Yamamoto T., Ueki T., Monden M., Kaji A.,
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of a newly identified member
RT of the cadherin family, PB-cadherin."
RT J. Biol. Chem. 271:11548-11556(1996).
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. PB-cadherins may have a role
CC in the morphological organization of pituitary gland and brain
CC tissues.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC Name=2;
CC IsoId=Q63315-1; Sequence=Displayed;
CC IsoId=Q63315-2; Sequence=VSP 000643, VSP 000644;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Strongly expressed in the pituitary gland and

```

the brain (in the inner granular and glomerular layers of the olfactory bulb, anterior olfactory nucleus, primary olfactory cortex, Purkinje cell layer of cerebellum, and pineal gland). Low expression in lung and heart. No expression in submandibular gland, thymus, liver, spleen, adrenal, and kidney.

CC -1- DEVELOPMENTAL STAGE: Expressed strongly in fetal brain.

CC -1- SIMILARITY: Contains 5 cadherin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D83348; BAA11894.1; -

CC EMBL; D83349; BAA11895.1; -

CC HSSP; P15116; INCU.

CC InterPro; IPR002126; Cadherin.

CC InterPro; IPR002233; Cadherin\_C\_term.

CC Pfam; PF01049; Cadherin\_5.

CC PRINTS; PR00205; CADHERIN.

CC SMART; SM00112; CA; 5.

CC PROSITE; PS00232; CADHERIN\_1; 2.

CC PROSITE; PS50268; CADHERIN\_2; 5.

CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat; Signal; Alternative splicing.

CC SIGNAL 1 33 POTENTIAL.

CC CHAIN 34 813 CADHERIN-22.

CC DOMAIN 33 621 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 622 642 POTENTIAL.

CC DOMAIN 643 813 CYTOPLASMIC (POTENTIAL).

CC DOMAIN 61 165 CADHERIN\_1.

CC DOMAIN 166 274 CADHERIN\_2.

CC DOMAIN 275 391 CADHERIN\_3.

CC DOMAIN 392 495 CADHERIN\_4.

CC DOMAIN 496 613 CADHERIN\_5.

CC CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC VARSPLIC 668 694 NDESGEODPEAYMSALRSYDFGEL -> TLGSGHGRGTS

CC /FtId=VSP\_000643.

CC /FtId=VSP\_000644.

CC VARSPLIC 695 813 Missing (in isoform 2).

CC SEQUENCE 813 AA; 87978 MW; 30BEA60B5DD2D467B CRC64;

CC Query March 48.9%; Score 45; DB 1; Length 813;

CC Best Local Similarity 54.5%; Pred. No. 21;

CC Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

CC

QY 2 WINWQHIDEH 12

Db 61 WYWNQFVEVE 71

RESULT 14

CAD\_HUMAN STANDARD; FR; 819 AA.

AC O66U0; O86U01; Q9N784;

DT 15-MAR-2004 (Rel. 43, last sequence update)

DT 15-MAR-2004 (Rel. 43, last annotation update)

DE Cadherin-24 precursor (UNC2834/PRO34009).

GN CDH24 OR CDH11.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI\_taxid=9606;

RP [1]

SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND INTERACTIONS WITH

CATENINS.

RX MEDLINE=22753805; PubMed=12734196;

RA Katsifas B.J., Nieman M.T., Wheelock M.J., Johnson K.R.;

RT "Characterization of cadherin-24, a novel alternatively spliced type II cadherin.";

RL J. Biol. Chem. 278:27513-27519 (2003).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE=22887296; PubMed=12975309;

RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J., Chen J., Chow B., Choi C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J., Grimaldi C., Gu Q., Hase P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J., Sesigiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A., Vandlen R., Watanabe C., Wiand D., Woods K., Xie M.-H., Yansura D., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I., Godowski P.;

RA "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";

RL Genome Res. 13:2265-2270(2003).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 3).

RA Blum H., Bauerach S., Mewes H.-W., Gassenhuber J., Wiemann S.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins. They preferentially interact with themselves in a homophilic manner in connecting cells; cadherins may thus contribute to the sorting of heterogeneous cell types. Cadherin-24 mediate strong cell-cell adhesion.

CC -1- SUBUNIT: Associates with alpha-, beta- and delta-catenins.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;

CC Name=1; Synonyms=Long form;

CC IsoId=Q86UP0-1; Sequence=Displayed;

CC Name=2; Synonyms=Short form;

CC IsoId=Q86UP0-2; Sequence=VSP\_008717;

CC Name=3;

CC IsoId=Q86UP0-3; Sequence=VSP\_008718, VSP\_008719;

CC Note=No experimental confirmation available;

CC -1- SIMILARITY: Contains 5 cadherin domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; AY260900; AAP20590.1; -

CC EMBL; AY260901; AAP20591.1; -

CC EMBL; AY358199; AA085866.1; -

CC EMBL; AL137477; CAB70758.1; -

CC PIR; T46418; T46418.

CC Gene; HGNC:14265; CDH24.

CC InterPro; IPR002126; Cadherin.

CC InterPro; IPR002233; Cadherin\_C\_term.

CC Pfam; PF00028; cadherin\_5.

CC PRINTS; PR00205; CADHERIN.

CC SMART; SM00112; CA; 4.

CC PROSITE; PS00232; CADHERIN\_1; 2.

CC PROSITE; PS50268; CADHERIN\_2; 5.

CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat; Signal; Multigene family; Alternative splicing.

CC SIGNAL 1 19 POTENTIAL.

CC PROPEP 21 44 POTENTIAL.

CC CHAIN 45 819 CADHERIN-24.

CC DOMAIN 45 641 EXTRACELLULAR (POTENTIAL).



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:30:55 ; Search time 31 Seconds  
(without alignments)  
152.670 Million cell updates/sec

Title: US-10-040-128-3  
Perfect score: 92  
Sequence: 1 DMINQMHIDEKNT 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	87	94.6	782	6	Q8NMW5	Q8NMW5 sus. scrofa
2	64	69.6	773	13	Q8AYD0	Q8AYD0 gallus galli
3	49	53.3	247	11	Q8C9N7	Q8C9N7 mus musculus
4	49	53.3	716	11	Q8C449	Q8C449 mus musculus
5	49	53.3	716	11	Q8C375	Q8C375 mus musculus
6	49	53.3	754	11	Q8BRK4	Q8BRK4 mus musculus
7	48	52.2	524	5	001321	001321 caenorhabditis
8	46	50.0	313	16	Q88HL3	Q88HL3 pseudomonas
9	45	48.9	176	10	Q8RYF1	Q8RYF1 avena strig
10	45	48.9	241	4	Q86TS8	Q86TS8 homo sapien
11	45	48.9	310	10	Q8W5G3	Q8W5G3 oryza sativ
12	45	48.9	333	2	Q9L3G9	Q9L3G9 erwinia rha
13	45	48.9	350	3	Q00870	Q00870 necrotia hae
14	45	48.9	493	4	Q96L07	Q96L07 homo sapien
15	45	48.9	566	3	Q9CA77	Q9CA77 cladosporiu
16	45	48.9	781	4	Q86UP1	Q86UP1 homo sapien

17	45	48.9	794	13	Q93264	Q93264 xenopus lae
18	45	48.9	796	4	Q96C29	Q96C29 homo sapien
19	45	48.9	796	11	Q8C7Q6	Q8C7Q6 mus musculus
20	45	48.9	819	4	Q86UP0	Q86UP0 homo sapien
21	45	48.9	839	5	Q9NR22	Q9NR22 strongyloce
22	44.5	48.4	963	10	Q9STU7	Q9STU7 arabidopsis
23	44	47.8	76	11	Q8BLT5	Q8BLT5 mus musculus
24	44	47.8	219	16	Q8BCK3	Q8BCK3 shewanella
25	44	47.8	271	5	Q8T9K7	Q8T9K7 drosophila
26	44	47.8	271	5	Q9VLX9	Q9VLX9 drosophila
27	44	47.8	337	11	Q8BLB5	Q8BLB5 mus musculus
28	44	47.8	340	11	Q8BLT4	Q8BLT4 mus musculus
29	44	47.8	500	16	Q9PT97	Q9PT97 straphylococ
30	44	47.8	500	16	Q8NMW1	Q8NMW1 straphylococ
31	44	47.8	551	13	Q8AMW2	Q8AMW2 gallus galli
32	44	47.8	630	4	Q8LY78	Q8LY78 homo sapien
33	44	47.8	785	11	Q8BMW2	Q8BMW2 mus musculus
34	44	47.8	790	4	Q8N522	Q8N522 homo sapien
35	44	47.8	790	13	Q91838	Q91838 xenopus lae
36	44	47.8	794	4	Q86UD2	Q86UD2 homo sapien
37	44	47.8	798	13	Q8QGH3	Q8QGH3 gallus galli
38	44	47.8	798	13	Q7ZYV7	Q7ZYV7 gallus galli
39	44	47.8	801	11	Q9Z0M3	Q9Z0M3 mus musculus
40	44	47.8	1356	5	Q8ICG0	Q8ICG0 plasmodium
41	44	47.8	1497	15	Q87G80	Q87G80 vibrio para
42	43	46.7	192	16	Q7VD46	Q7VD46 prochloroco
43	43	46.7	246	16	Q7VZB4	Q7VZB4 prochloroco
44	43	46.7	248	16	Q7VA86	Q7VA86 prochloroco
45	43	46.7	253	16	Q7U5H3	Q7U5H3 synechococc

## ALIGNMENTS

RESULT 1  
ID Q8NMW5 PRELIMINARY: PRT: 782 AA.  
AC Q8NMW5;  
DT 01-MAR-2002 (TRENBLREL. 20, Created)  
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
DE VE cadherin.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
OC NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hirano M., Hirano K., Nishimura J., Kanade H.;  
RT "Transcriptional up-regulation of p27kip1 during contact-induced  
RT growth arrest in the endothelial cells."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99119339; PubMed=9918866;  
RA Hirano M., Nitro N., Hirano K., Nishimura J., Hartshorne D.J.,  
RA Kanade H.;  
RT "Expression, subcellular localization and cloning of the 130 kDa  
RT regulatory subunit of myosin phosphatase in porcine aortic endothelial  
RT cells."  
RL Biochem. Biophys. Res. Commun. 254:490-496 (1999).  
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
CC THEY DIFFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
CC SORTING OF HETEROGENEOUS CELL TYPES (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
DR EMBL: AB046120; BAB82983.1; -;  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005509; F:calcium ion binding; IEA.  
DR GO: GO:0005194; F:cell adhesion molecule activity; IEA.  
DR GO: GO:0007155; P:cell adhesion; IEA.  
DR GO: GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro: IPR002126; Cadherin.

DR InterPro; IPR000233; Cadherin\_C\_term.  
 DR Pfam; PF00028; cadherin\_5.  
 DR Pfam; PF01049; Cadherin\_C\_term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS0268; CADHERIN\_2; 5.  
 DR KEGG; Calcium-binding; Cell adhesion; Glycoprotein.  
 KW Calcium; 782 AA; 87536 MW; 681D12D017CBI083 CRC64;  
 SQ SEQUENCE

Query Match 94.6%; Score 87; DB 6; Length 782;  
 Best local Similarity 100.0%; Pred. No. 1.1e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DWIMNQMHIIDEK 14  
 Db 45 DWIMNQMHIIDEKN 58

RESULT 2  
 Q8AYD0 PRELIMINARY; PRT; 773 AA.

AC Q8AYD0;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Vascular endothelial cadherin precursor.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Romero S., Jaffredo T., Dunon D.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF522067; AAN33002.1;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR000233; Cadherin\_C\_term.  
 DR Pfam; PF00028; cadherin\_5.  
 DR Pfam; PF01049; Cadherin\_C\_term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS0268; CADHERIN\_2; 5.  
 DR KEGG; Signal.  
 FT CHAIN 1 18 POTENTIAL.  
 FT SIGNAL 43 773 VASCULAR ENDOTHELIAL CADHERIN.  
 FT SEQUENCE 773 AA; 86842 MW; ACS9CE1119043F4 CRC64;

Query Match 69.6%; Score 64; DB 13; Length 773;  
 Best local Similarity 83.3%; Pred. No. 0.064;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DWIMNQMHIIDE 12  
 Db 41 DWIMNQMHIIDE 52

RESULT 3  
 Q8C9N7 PRELIMINARY; PRT; 247 AA.

AC Q8C9N7;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Cadherin 8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK041720; BAC31041.1;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
 DR InterPro; IPR002126; Cadherin.  
 DR Pfam; PF00028; cadherin\_2.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 1.  
 DR PROSITE; PS00232; CADHERIN\_1; 1.  
 DR PROSITE; PS0268; CADHERIN\_2; 2.  
 SQ SEQUENCE 247 AA; 27565 MW; DF7551A97A29289D CRC64;

Query Match 53.3%; Score 49; DB 11; Length 247;  
 Best local Similarity 63.6%; Pred. No. 5.4;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WIMNQMHIIDE 12  
 Db 63 WIMNQMHIIDE 73

RESULT 4  
 Q8C449 PRELIMINARY; PRT; 716 AA.

AC Q8C449;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Cadherin 8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Hippocampus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK083092; BAC38758.1;  
 DR MGI; MGI:107434; Cdh8.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR00235; Cadherin\_C\_term.  
 DR Pfam; PF00028; cadherin\_5.  
 DR Pfam; PF01049; Cadherin\_C\_term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS0268; CADHERIN\_2; 5.  
 SQ SEQUENCE 716 AA; 79145 MW; 976FF5D845F938BD CRC64;

Query Match 53.3%; Score 49; DB 11; Length 716;  
 Best local Similarity 63.6%; Pred. No. 17;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WIMNQMHIIDE 12



```

Db      63 WWMQMFVLEE 73

RESULT 5
Q8C375      PRELIMINARY;      PRT;      716 AA.
AC      Q8C375;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Cadherin 8.
GN      CDH8.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573 (2002).
DR      EMBL; AK086711; BAC39724.1; -.
DR      MGD; MGI:107434; Cdh8.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005509; F:calcium ion binding; IEA.
DR      GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR      InterPro; IPR002126; Cadherin.
DR      InterPro; IPR000233; Cadherin_C_term.
DR      Pfam; PF00028; cadherin; 5.
DR      Pfam; PF01049; Cadherin_C_term; 1.
DR      PRINTS; PR00205; CADHERIN.
DR      SMART; SM00112; CA; 5.
DR      PROSITE; PS00232; CADHERIN_1; 3.
DR      PROSITE; PS00268; CADHERIN_2; 5.
DR      PROSITE; PS00268; CADHERIN_2; 5.
SQ      SEQUENCE 716 AA; 79149 MW; 5BBD598F638624A CRC64;

Query Match      53.3%; Score 49; DB 11; Length 716;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 WWMQMHIDE 12
      |:|:|:|:|
Db      63 WWMQMFVLEE 73

RESULT 6
Q8BRK4      PRELIMINARY;      PRT;      754 AA.
AC      Q8BRK4;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Cadherin 8.
GN      Cadherin 8.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Cortex;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573 (2002).
DR      EMBL; AK04046; BAC31751.1; -.
DR      EMBL; AK04046; BAC31751.1; -.
DR      GO; GO:0016020; C:membrane; IEA.

```

```

DR      GO; GO:0005509; F:calcium ion binding; IEA.
DR      GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR      InterPro; IPR002126; Cadherin.
DR      InterPro; IPR000233; Cadherin_C_term.
DR      Pfam; PF00028; cadherin; 5.
DR      Pfam; PF01049; Cadherin_C_term; 1.
DR      PRINTS; PR00205; CADHERIN.
DR      SMART; SM00112; CA; 5.
DR      PROSITE; PS00232; CADHERIN_1; 3.
DR      PROSITE; PS00268; CADHERIN_2; 5.
DR      PROSITE; PS00268; CADHERIN_2; 5.
SQ      SEQUENCE 754 AA; 83624 MW; 149C3C70D3F21568 CRC64;

Query Match      53.3%; Score 49; DB 11; Length 754;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 WWMQMHIDE 12
      |:|:|:|:|
Db      63 WWMQMFVLEE 73

RESULT 7
001321      PRELIMINARY;      PRT;      524 AA.
AC      001321; P91858;
DT      01-JUL-1997 (TrEMBLrel. 04, Created)
DT      01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      P26H9.1 protein.
GN      P26H9.1.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC      Rhabditidae; Pelodierinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Baynes C.;
RL      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA      Boulfield J., Burton V., Connell M., Copsey T., Cooper J., Coulson A.,
RA      Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA      Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA      Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA      Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA      Parsons V., Percy C., Ritken L., Roopra A., Saunders D., Showkeen R.,
RA      Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA      Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA      Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans.";
RL      Nature 368:32-38 (1994).
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Sims M.;
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Z61516; CAB04206.1; -.
DR      EMBL; Z75535; CAB04206.1; JOINED.
DR      EMBL; Z75535; CAB99829.1; -.
DR      EMBL; Z81516; CAA99829.1; JOINED.
DR      PIR; T20872; T20872.
DR      WormPep; F26H9.1; CE09706.
DR      InterPro; IPR008945; Skp1_Skp2.
SQ      SEQUENCE 524 AA; 60462 MW; 6A85B0C0DCDD2B76 CRC64;

Query Match      52.2%; Score 48; DB 5; Length 524;
Best Local Similarity 53.8%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 WWMQMHIDEKN 14
      |:|:|:|:|

```

Db 486 WDMHORNIDEHRN 498

RESULT 8

Q88HL3 PRELIMINARY; PRT; 313 AA.

AC Q88HL3; 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Nickel ABC transporter, permease protein.

GN NIKB OR PP3343.

OS Pseudomonas putida (strain KT2440).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI\_TaxID=160488;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22423060; PubMed=12534463;

RA Nelson K.E., Meinel C., Paulsen I.T., Dodson R.J., Hilbert H., Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M., Brinkac L., Bauman M., DeBoy R.T., Daugherty S., Kolonay J., Madupu R., Nelson W., White O., Peterson J., Kouri H., Hance I., Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazed A., Utecherbeck T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H., Lauber J., Stepanic D., Hoheisel J., Straetz M., Helm S., Kiewitz C., Eissen J., Timmis K.N., Duesterhoeft A., Tsemmler B., Fraser C.M.;

RA "Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440.";

RL EMBL; Microbiol. 4:799-808(2002).

DR EMBL; AB016786; AAN68947.1; -.

DR TIGR; PP3343; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005215; P:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000515; BPD\_transp.

DR Pfam; PF00528; BPD\_transp; 1.

KW Complete proteome.

SQ SEQUENCE 313 AA; 34571 MW; 0480D2B8E523F87C CRC64;

QY 2 WDMHORNIDEHRN 10

Db 71 WDMHORNIDEHRN 79

Query Match 50.0%; Score 46; DB 16; Length 313;

Best Local Similarity 55.6%; Pred. No. 22;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 9

Q88YFL PRELIMINARY; PRT; 176 AA.

AC Q88YFL; 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Putative resistance protein (Fragment).

GN RGA.

OS Avena strigosa (black oat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;

OC Avenae; Avena.

OX NCBI\_TaxID=38783;

RN [1]

RP SEQUENCE FROM N.A.

RA Loarce Y., Irigoyen M.L., Fominaya A., Ferrer E.;

RT "Resistance gene analogous in Avena strigosa.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ437574; CAD26862.1; -.

DR GO; GO:0005524; P:ATP binding; IEA.

DR GO; GO:0006915; P:apoptosis; IEA.

DR InterPro; IPR002182; NB-ARC.

DR Pfam; PF00931; NB-ARC; 1.

FT NON TER 1

FT NON TER 176

SQ SEQUENCE 176 AA; 20133 MW; D60982A15C6FE2F6 CRC64;

Query Match 48.9%; Score 45; DB 10; Length 176;

Best Local Similarity 58.3%; Pred. No. 17;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDMHORNIDEHRN 14

Db 62 WDMHORNIDEHRN 73

RESULT 10

Q86TS8 PRELIMINARY; PRT; 241 AA.

AC Q86TS8; 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Human full-length cDNA 5-PRIME end of clone CSDB0107P19 of placenta of Homo sapiens (Human).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Genoscope;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Li W.B., Gruber C., Jessee J., Polayes D.;

RT "Full-length cDNA libraries and normalization.";

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX248750; CAD6557.1; -.

DR GO; GO:0046821; C:extrachromosomal DNA; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0007156; P:homophilic cell adhesion; IEA.

DR InterPro; IPR002126; Cadherin.

DR Pfam; PF00028; cadherin; 2.

DR PRINTS; PR00205; CADHERIN.

DR SMART; SM00112; CA; 2.

DR PROSITE; PS00232; CADHERIN\_1; 1.

DR PROSITE; PS50268; CADHERIN\_2; 2.

KW Plasmid.

FT NON TER 241

SQ SEQUENCE 241 AA; 26348 MW; 4649831B55424604 CRC64;

QY 2 WDMHORNIDEHRN 12

Db 46 WDMHORNIDEHRN 56

Query Match 48.9%; Score 45; DB 4; Length 241;

Best Local Similarity 54.5%; Pred. No. 24;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 11

Q8W5G3 PRELIMINARY; PRT; 310 AA.

AC Q8W5G3; 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Putative transcription factor.

GN OSUNBA002324.23.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae;  
OC Erihartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxId=4530;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Buell C.R., Yuan Q., Guyang S., Liu J., Moffatt K.S., Hill J.N.,  
RA Gansberger K., Brenner M., Birgeess S., Hance M., Shvatsitsyn M.,  
RA Tseltrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,  
RA Vanden S.E., Uteclback T.R., Feldblum T.V., Kalb E., Quackenbush J.,  
RA Salzberg S.L., White O., Frazer C.M.,  
RT "Oryza sativa chromosome 3 BAC OSUHBa002J24 genomic sequence.",  
RL submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS.  
DR EMBL; AC090713; ALU34136.1; -.  
DR Gramene; QBW5G3; -.  
DR InterPro; IPR001092; HLH\_basic.  
DR Pfam; PF00010; HLH; 1.  
DR SMART; SMO0353; HLH; 1.  
DR PROSITE; PSS0888; HLH\_2; 1.  
SQ SEQUENCE 310 AA; 34267 MW; FCB99D4910830A1D CRC64;

Query Match	48.9%	Score 45;	DB 10;	Length 310;
Best Local Similarity	50.0%	Pred. No. 32;		
Matches	7; Conservative	2; Mismatches	5; Indels	0; Gaps
QY	1 DWINQMHIDPEKN 14			
	:			
	:			
Db	4 DWIERRRREERN 17			

QY	2	WINNOMIDEEN	14
Db	75	WAMTQTHSNPK	877
RESULT 13			
000870			
ID	Q00870	PRELIMINARY;	
AC	Q00870;		
DT	01-NOV-1996 (TREMblrel)	0	
DT	01-NOV-1996 (TREMblrel)	0	
DT	01-MAR-2002 (TREMblrel)	2	
DE	Keystone hydrates precu		
GN	XIS.		
OS	Nectria haematococca.		
OC	Eukaryota; Fungi; Ascomyc		
OC	Hypocremycetidae; Hypocr		
OX	NCBI_TaxID=140110;		
RN	[1]_		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=FBI-S; TISSUE=Mycel		
XC	MEDLINE=95385711; PubMed=		
RA	Ia D.; Chung K.R.; Smith h		
RT	"The Fusarium solani gene		
RT	enzyme that catalyzes dete		
RL	Mol. Plant Microbe Interact		
KW	EMBL; U39639; AAA887627.1;		
FT	SIGNAL.		
FT	SIGNAL.	2	19
FT	CHAIN	10	360

FT	CHAIN	20	350	KIEVITONE HYDRATASE.
SQ	SEQUENCE	350 AA;	38970 MW;	CE09B07A7B34C91B CRC64;

RESULT 12	
09L3G9	
PRELIMINARY;	PRT; 333 AA.
ID 09L3G9	
AC 09L3G9	
DT 01-OCT-2000 (TEMBLrel. 15, Created)	
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)	
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)	
DE Endoglucanase precursor (EC 3.2.1.4).	
EN CELA.	

Query Match	48.9%	Score 45;	DB 3;	Length 350;
Best Local Similarity	63.6%;	Pred No. 36;		
Matches	7;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0.
QY	1 DWIMNQHIDE 11			
:	: 0       :			
-	- - - - -			
Db	219 NWTWFOHIEP 229			

OS *Erwinia raphanocritici*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Erythrobacteriales;  
OC Enterobacteriaceae; *Erwinia*.  
OX NCBI\_TaxID=55212;  
RN [1]\_TaxID=55212;  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCPP2989;  
RA Saarijahi H.T.;  
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RI

NS  
12;  
RP  
RC  
STRAIN=NCPE2989;  
RA  
RIEKKI R.;  
RT  
RT  
possesses genes homologous to the type II secretion pathway";  
RL  
Submitted (May-2000) to the EMBL/Genbank/DBJ databases.  
EMBL: A37623f8 - CYPB0603.

DR	GO: 0020636;	Can63803.1;	-
DR	GO: 0008810;	F:cellulase activity;	IEA.
GO	GO: 0005975;	P:carbohydrate metabolism;	IEA.
DR	InterPro: IPR02037;	Glyco hydro 8.	
DR	InterPro: IPR008928;	Glyco trans_6np.	
DR	Pfam: PF01270;	Glyco hydro 8; 1.	
DR	PRINTS: PR00735;	GHYDRASE8.	
KW	Glycosidase; Hydrolase;	Signal.	
FT	SIGNAL	1	23
FT			POTENTIAL.
FT	CHAIN	24	333
SEQUENCE	333 AA;	37784 MM;	ENDOCYCLICANASE.
FT			CRG64:

Query Match 48.9%; Score 45; DB 2; Length 333;  
Best Local Similarity 46.2%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; G

RESULT	13		
000870			
ID	Q00870	PRELIMINARY;	PRT; 350 AA.
AC	Q00870;		
DT	01-NOV-1996 (TREMblrel. 01, Created)		
DT	01-NOV-1996 (TREMblrel. 01, last sequence update)		
DT	01-MAR-2002 (TREMblrel. 20, last annotation update)		
DE	Kievitone hydratase precursor.		
GN	KHS.		
OS	Nectria haematococcica.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes		
CC	Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.		
CC	NCBI_TaxID=140110;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=FBI-S; TISSUE=Mycelium;		
RX	MEDLINE=95183711; PubMed=7655061.		

RA Li D., Chung K. R., Smith D. A., Schardl C. L.;  
 RT "the Fusarium solani gene encoding Kievitone hydratase, a secreted  
 enzyme that catalyzes detoxification of a bean phytoalexin."; *Mol. Plant Microbe Interact.* 8:386-397(1995).  
 DR EMBL: L33639, AAA87627.1; --  
 KW SIGNAL..  
 FT SIGNAL..  
 FT CHAIN 1 19 KIEVITONE HYDRATASE.  
 FT 20 350  
 SQ SEQUENCE 350 AA; 38970 MW; CE05B07A7B34C91B CRC64;

Query Match	48.9%	Score 45	DB 3	Length 350
Best Local Similarity	63.6%	Pred. No. 36		
Matches	7	Conservative	1	Mismatches 3
				Indels 0
				Gaps 0
QY	1	DWIMNQHIDE	11	
	:	0		
Db	219	NWTFOWHIE	229	

RESULT 14		
Q96L07		
ID	Q96L07	PRELIMINARY;
ID	Q96L07	PRT; 493 AA.
DT	01-DEC-2001	(TRIMBLrel. 19, Created)
DT	01-DEC-2001	(TRIMBLrel. 19, Last sequence update)
DT	01-OCT-2003	(TRIMBLrel. 25, Last annotation update)
DE	Hypochemical protein FLJ25193.	
OS	Homo sapiens (Human).	

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thyroid;  
RA Nishimura K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira  
AC

RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kanibara  
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,  
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,  
RA Irie K., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami  
RA Nagai K., Isogai T., Sugano S.;  
RA "NEO human cDNA sequencing project.";  
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ  
DR EMBL, AK057922; BAB71613.1; -  
DR GO: GO:0016020; F:calcium ion binding, IEA.  
DR GO: GO:0005509; F:calcium ion binding, IEA.

DR GO; GO:0007156, P:homophilic cell adhesion; IEA.  
DR GO; GO:0007155, P:cell adhesion; IEA.  
DR GO; GO:0005134, P:cell adhesion molecule activity; IEA.  
DR GO; GO:0005134, P:cell adhesion molecule activity; IEA.  
DR Pfam; PF000028, cadherin; 4.  
DR InterPro; IPR002126, Cadherin.

DR SMART; SM00112; CA; 4.  
 DR PROSITE; PS00232; CADHERIN\_1; 2.  
 DR PROSITE; PS00268; CADHERIN\_2; 4.  
 KW Hypothetical protein; Calcium; Calcium-binding; Cell adhesion;  
 KW Glycoprotein.  
 SQ SEQUENCE 493 AA; 53618 MW; 33F10DF63AF09C1E CRC64;

Query Match 48.9%; Score 45; DB 4; Length 493;  
 Best Local Similarity 54.5%; Pred. No. 52;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 DWIMNQHIDE 12  
 ||:||||: ||  
 Db 46 DWIMNQHIDE 56

## RESULT 15

O9C477 PRELIMINARY; PRT; 566 AA.  
 AC O9C477;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE CATI catalase (EC 1.11.1.6).  
 GN CAT-1.  
 OS Cladosporium fulvum (Fulvia fulva).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;  
 OC Dothideomycetes et Chaetochytriomycetes incertae sedis;  
 OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.  
 OX NCBI TaxID=5499;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Race 4;  
 RX MEDLINE=20570112; PubMed=11121097;  
 RA Bussink H.-J.;  
 RT "Identification of two highly divergent catalase genes in the fungal  
 RT tomato pathogen, Cladosporium fulvum.";  
 RL Eur. J. Biochem. 268:15-24(2001).  
 CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC EMBL; AF222055; AAC53518.1; -.  
 DR HSSP; P04040; 1F4J.  
 DR GO; GO:0004096; F:catalase activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0004601; F:peroxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0006979; P:response to oxidative stress; IEA.  
 DR InterPro: IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR PRODOM; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 566 AA; 63896 MW; 59023AA3FEFF32334 CRC64;

Query Match 48.9%; Score 45; DB 3; Length 566;  
 Best Local Similarity 40.0%; Pred. No. 61;  
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 DWIMNQHIDEKNT 15  
 ||:||||: ||  
 Db 238 DWIMNQHIDEKNT 252

Search completed: July 29, 2004, 11:37:09  
 Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:30:30 ; Search time 6.6667 Seconds

(without alignments)  
117.158 Million cell updates/sec

Title: US-10-040-128-3

Perfect score: 92

Sequence: 1 DWINQMTHDEKNT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	784	1	CAD5_HUMAN P33151 homo sapien
2	87	94.6	782	1	CAD5_PIG O02840 sus scrofa
3	87	94.6	784	1	CAD5_MOUSE P55284 mus musculu
4	57	62.0	772	1	CAD1_HUMAN Q9H159 homo sapien
5	49	53.3	799	1	CAD8_HUMAN P55286 homo sapien
6	49	53.3	799	1	CAD8_MOUSE P97291 mus musculu
7	49	53.3	799	1	CAD8_RAT O54800 rattus norv
8	45	48.9	792	1	CAD8_CHICK O93319 gallus gall
9	45	48.9	796	1	CAD8_HUMAN P55287 homo sapien
10	45	48.9	796	1	CAD8_MOUSE P55288 mus musculu
11	45	48.9	813	1	CAD4_MOUSE O947P5 mus musculu
12	45	48.9	813	1	CAD4_RAT O63315 rattus norv
13	45	48.9	819	1	CAD0_HUMAN O86up0 homo sapien
14	45	48.9	828	1	CAD0_HUMAN Q9Uj99 homo sapien
15	44	47.8	785	1	CAD7_CHICK Q90763 gallus gall
16	44	47.8	785	1	CAD7_HUMAN Q9U1b5 homo sapien
17	44	47.8	790	1	CAD1_HUMAN Q13634 homo sapien
18	44	47.8	794	1	CAD0_HUMAN P55289 homo sapien
19	44	47.8	801	1	CAD0_HUMAN P57201 buchnera ap
20	43	46.7	835	1	HISX_HUMAN O09671 escherichia
21	42.5	46.2	388	1	YIBA_BACSU O03086 bacillus su
22	42	45.7	427	1	CISY_SALTY O68883 salmonella
23	42	45.7	583	1	GHS2_CAEHL Q93650 caenorhabdi
24	41.5	45.1	392	1	OYEB_SCHPO Q09671 escherichia
25	41	44.6	314	1	TRPG_ASFAM P18483 aspergillus
26	41	44.6	768	1	CAD4_HUMAN O94688 homo sapien
27	41	44.6	788	1	CAD4_HUMAN P55280 rattus norv
28	41	44.6	789	1	CAD9_HUMAN Q9U1b4 homo sapien
29	41	44.6	789	1	CAD9_CHICK P79395 gallus gall
30	41	44.6	790	1	CAD6_CHICK P55285 homo sapien
31	41	44.6	790	1	CAD6_HUMAN P55285 homo sapien
32	41	44.6	790	1	CAD6_HUMAN P55285 homo sapien
33	40.5	44.0	44	1	OP01_OPICA P83313 opisthophtha

34	40.5	44.0	44	1	OP02_OPICA P83314 opisthophtha
35	40.5	44.0	757	1	IKKX_MOUSE O88351 mus musculu
36	40.5	44.0	757	1	IKKX_RAT O9q778 rattus norv
37	40	43.5	146	1	HBBC_HOPLI P82316 hoplosternu
38	40	43.5	200	1	LEUD_HAEIN P44438 haemophilus
39	40	43.5	373	1	LYSI_YEAST P38998 saccharomyc
40	40	43.5	433	1	UROK_MOUSE P06869 mus musculu
41	40	43.5	475	1	PEM2_ARATH O94480 arabidopsis
42	40	43.5	508	1	TNSD_ECOLI P13921 escherichia
43	40	43.5	518	1	NIRK_THIFE P15052 thiodocellu
44	40	43.5	1742	1	GUNA_CALSA P22534 caldocellum
45	40	43.5	1812	1	SUIS_SUNMU O62653 sunus muri

## ALIGNMENTS

RESULT 1  
CAD5\_HUMAN STANDARD; PRT; 784 AA.  
ID CAD5\_HUMAN P33151;  
AC P33151;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5)  
DN (784 antigen) (CD144 antigen).  
GN CDH5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Endothelial cells;  
RX MEDLINE=95353875; PubMed=7627717;  
RA Breviario F., Cavada L., Corada M., Martin-Padura I., Navarro P.,  
RA Goley J., Introna M., Gullino D., Lampugnani M.G., Dejana F.;  
RT "Functional properties of human vascular endothelial cadherin  
RT (784/cadherin-5), an endothelium-specific cadherin";  
RT Arterioscler. Thromb. Vasc. Biol. 15:1229-1239(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=97362755; PubMed=9219219;  
RA Ali J., Liao F., Martens B., Muller W.A.;  
RT "Vascular endothelial cadherin (VE-cadherin): cloning and role in  
RT endothelial cell-cell adhesion";  
RT Microcirculation 4:267-277(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC PubMed=10861224;  
RX Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;  
RA "Identification of three human type-II classic cadherins and frequent  
RT heterophilic interactions between different subclasses of type-II  
RT classic cadherins";  
RL Biochem. J. 349:159-167(2000).  
RN [4]  
RP SEQUENCE OF 5-784 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91283540; PubMed=2059658;  
RA Suzuki S., Sano K., Tanihara H.;  
RT "Diversity of the cadherin family: evidence for eight new cadherins  
RT in nervous tissue";  
RL Cell Regul. 2:261-270(1991).  
RN [5]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Endothelial cells;  
RX MEDLINE=92394977; PubMed=1522121;  
RA Lampugnani M.G., Resnati M., Raiteri M., Pigott R., Pisacane A.,  
RA Hogen G., Ruco L.P., Dejana F.;  
RT "A novel endothelial-specific membrane protein is a marker of  
RT cell-cell contacts";  
RL J. Cell Biol. 118:1511-1522(1992).

```

CC CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC CC They preferentially interact with themselves in a homophilic
CC CC manner in connecting cells; cadherins may thus contribute to the
CC CC sorting of heterogeneous cell types. This cadherin may play a
CC CC important role in endothelial cell biology through control of the
CC CC cohesion and organization of the intercellular junctions. It
CC CC associates with alpha-catenin forming a link to the cytoskeleton.
CC CC SUBCELLULAR LOCATION: Type I membrane protein. Found at cell-cell
CC CC boundaries and probably at cell-matrix boundaries.
CC CC -!- TISSUE SPECIFICITY: Endothelial tissues and brain.
CC CC -!- SIMILARITY: Contains 5 cadherin domains.
CC CC -!- DATABASE: NAME=PRO; NOTE=CD guide CD144 entry;
CC CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd144.htm".
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: X79981; CA56306.1; -.
CC CC DR EMBL: U84722; AAB41796.1; -.
CC CC DR EMBL: AB035304; BAA87418.1; -.
CC CC DR EMBL: X59796; CAA42468.1; -.
CC CC DR PIR: S49893; ICHUC5.
CC CC DR HSPSP: P15116; INCU.
CC CC DR GlycoSiteDB: P3151; -.
CC CC DR Genew: HGNC:1764; CDH5.
CC CC -----
CC CC MIM: 601120; -.
CC CC DR GO: 0005624; C:membrane fraction; TAS.
CC CC DR GO: 0005886; C:plasma membrane; TAS.
CC CC DR GO: 0007156; P:homophilic cell adhesion; TAS.
CC CC DR InterPro: IPR002126; Cadherin.
CC CC DR InterPro: IPR000233; Cadherin_C_term.
CC CC DR Pfam: PF00028; cadherin; 5.
CC CC DR Pfam: PF01049; Cadherin_C_term; 1.
CC CC DR PRINTS: PR00205; CADHERIN.
CC CC DR SMART: SM00112; CA; 5.
CC CC DR PROSITE: PS00232; CADHERIN_1; 3.
CC CC DR PROSITE: PS50268; CADHERIN_2; 5.
CC CC DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
CC CC KW Signal.
CC CC FT SIGNAL 1 25 POTENTIAL.
CC CC FT PROPEP 26 47 POTENTIAL.
CC CC FT CHAIN 48 784 VASCULAR ENDOTHELIAL-CADHERIN.
CC CC FT DOMAIN 48 593 EXTRACELLULAR (POTENTIAL).
CC CC FT TRANSMEM 594 620 POTENTIAL.
CC CC FT DOMAIN 621 784 CYTOPLASMIC (POTENTIAL).
CC CC FT DOMAIN 48 151 CADHERIN 1.
CC CC FT DOMAIN 152 258 CADHERIN 2.
CC CC FT DOMAIN 259 372 CADHERIN 3.
CC CC FT DOMAIN 373 477 CADHERIN 4.
CC CC FT DOMAIN 478 593 CADHERIN 5.
CC CC FT DOMAIN 736 753 SER-RICH.
CC CC FT CARBOHYD 61 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 517 517 T -> I (IN REF. 2 AND 4).
CC CC FT CONFLICT 517 517
CC CC FT SEQUENCE 784 AA; 87516 MW; F643BFC22A599DE CRC64;
CC CC -----

```

```

CC CC -----
CC CC RESULT 2
CC CC ID CAD5_PIG STANDARD; PRT; 782 AA.
CC CC AC 002840;
CC CC DT 15-JUL-1998 (Rel. 36, Created)
CC CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC CC DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5).
CC CC GN CDH5.
CC CC OS Sus scrofa (Pig).
CC CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC CC OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
CC CC OX NCBI_TaxID=9823;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RA Kishaw P.J.;
CC CC RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC CC CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC CC They preferentially interact with themselves in a homophilic
CC CC manner in connecting cells; cadherins may thus contribute to the
CC CC sorting of heterogeneous cell types. This cadherin may play a
CC CC important role in endothelial cell biology through control of the
CC CC cohesion and organization of the intercellular junctions. It
CC CC associates with alpha-catenin forming a link to the cytoskeleton
CC CC (By similarity).
CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Found at cell-cell
CC CC boundaries and probably at cell-matrix boundaries (By similarity).
CC CC -!- SIMILARITY: Contains 5 cadherin domains.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: Y13919; CAA74225.1; -.
CC CC DR HSPSP: P09803; IBDH.
CC CC DR InterPro: IPR002126; Cadherin.
CC CC DR InterPro: IPR000233; Cadherin_C_term.
CC CC DR Pfam: PF00028; cadherin; 5.
CC CC DR Pfam: PF01049; Cadherin_C_term; 1.
CC CC DR PRINTS: PR00205; CADHERIN.
CC CC DR SMART: SM00112; CA; 5.
CC CC DR PROSITE: PS00232; CADHERIN_1; 3.
CC CC DR PROSITE: PS50268; CADHERIN_2; 5.
CC CC DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
CC CC KW Signal.
CC CC FT SIGNAL 1 22 POTENTIAL.
CC CC FT PROPEP 23 44 POTENTIAL.
CC CC FT CHAIN 45 782 VASCULAR ENDOTHELIAL-CADHERIN.
CC CC FT DOMAIN 45 592 EXTRACELLULAR (POTENTIAL).
CC CC FT TRANSMEM 593 619 POTENTIAL.
CC CC FT DOMAIN 620 782 CYTOPLASMIC (POTENTIAL).
CC CC FT DOMAIN 45 148 CADHERIN 1.
CC CC FT DOMAIN 149 255 CADHERIN 2.
CC CC FT DOMAIN 256 370 CADHERIN 3.
CC CC FT DOMAIN 371 475 CADHERIN 4.
CC CC FT DOMAIN 476 592 CADHERIN 5.
CC CC FT DOMAIN 736 751 SER-RICH.
CC CC FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT SEQUENCE 782 AA; 87546 MW; 7403F974E2DF82F CRC64;
CC CC -----

```

```

QY 1 DWIMQMIDEKNT 15
DB 48 DWIMQMIDEKNT 62
Query Match 100.0%; Score 92; DB 1; Length 784;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 94.6%; Score 87; DB 1; Length 782;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;

```

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKN 14  
 Db 45 DWIMQMHIDEKN 58

RESULT 3  
 CAD5\_MOUSE  
 ID CAD5\_MOUSE STANDARD; PRT; 784 AA.  
 AC P55284; O35542;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5).  
 GN CDH5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain capillary;  
 RX MEDLINE=96141083; PubMed=8555485;  
 RA Breier G., Breviaro F., Cavada L., Berthier R., Schnerch H.,  
 RA Gotsch U., Westweber D., Rissau W., Dejana E.,  
 RT "Molecular cloning and expression of murine vascular endothelial-  
 RT cadherin in early stage development of cardiovascular system.";  
 RL Blood 87:630-641(1996).  
 RN [2]  
 RP REVISIONS TO 67-70.  
 RA Breviaro F.;  
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RC STRAIN=BALB/c; TISSUE=Breast carcinoma;  
 RX MEDLINE=97364256; PubMed=9220534;  
 RA Matuyoshi N., Toda K.-I., Horiguchi Y., Tanaka T., Nakagawa S.,  
 RA Takeichi M., Imanura S.,  
 RT "In vivo evidence of the critical role of cadherin-5 in murine  
 RT vascular integrity".  
 RL Proc. Assoc. Am. Physicians 109:362-371(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=1247932;  
 RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalov D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences".  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. This cadherin may play a  
 CC important role in endothelial cell biology through control of the  
 CC cohesion and organization of the intercellular junctions.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X83930; CA58782.2; -;  
 DR EMBL; D63942; BAA22617.1; -;  
 DR EMBL; BC054790; AAH54790.1; -;  
 DR HSSP; P15116; INCU.  
 DR MGD; MG1:105057; Cdh5.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR002233; Cadherin\_C\_term.  
 DR Pfam; PF00208; cadherin; 5.  
 DR Pfam; PF01049; Cadherin\_C\_term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS50268; CADHERIN\_2; 5.  
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 24  
 FT PROPEP 25 45  
 FT CHAIN 46 784  
 FT DOMAIN 46 593  
 FT TRANSMEM 594 620  
 FT DOMAIN 621 784  
 FT DOMAIN 46 149  
 FT DOMAIN 150 257  
 FT DOMAIN 257 371  
 FT DOMAIN 372 476  
 FT DOMAIN 477 593  
 FT DOMAIN 728 753  
 FT CARBOHYD 59 59  
 FT CARBOHYD 155 155  
 FT CARBOHYD 441 441  
 FT CARBOHYD 523 523  
 FT CARBOHYD 535 535  
 SQ SEQUENCE 784 AA; 87902 MW; 7B75698DE2F7E160 CRC64;

Query Match 94.6%; Score 87; DB 1; Length 784;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKN 14  
 Db 46 DWIMQMHIDEKN 59

RESULT 4  
 CADJ\_HUMAN  
 ID CADJ\_HUMAN STANDARD; PRT; 772 AA.  
 AC Q9H159;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Cadherin-19 precursor (UNC478/PRO941).  
 GN CDH19 OR CDH7L2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20453188; PubMed=10995570;  
 RX Kooles P., Van Imesch G., van Roy F.,  
 RA "Characterization of three novel human cadherin genes (CDH7, CDH19,  
 RT and CDH20) clustered on chromosome 18q22-q23 and with high homology  
 RT to chicken cadherin-7".

```

RL Genomics 68:283-295(2000).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in many tissues, with the exception
CC of uterus.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ007607; CAC13126.1; -.
DR EMBL; AY358654; AA089017.1; -.
DR HSSP; P15116; INCU.
DR Genew; HGNC:1758; CDH19.
DR MIM; 603016; -.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0007156; P:homophilic cell adhesion; NAS.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; Cadherin_C_term; 1.
DR SMART; SMO0112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS50268; CADHERIN_2; 5.
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 43 POTENTIAL.
FT CHAIN 44 772 CADHERIN-19.
FT DOMAIN 44 596 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 597 617 POTENTIAL.
FT DOMAIN 618 772 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 44 148 CADHERIN_1.
FT DOMAIN 149 256 CADHERIN_2.
FT DOMAIN 257 370 CADHERIN_3.
FT DOMAIN 371 470 CADHERIN_4.
FT DOMAIN 470 581 CADHERIN_5.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 508 508 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 516 516 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 772 AA; 87002 MW; 650AD27480343C39 CRC64;

Query Match 62.0%; Score 57; DB 1; Length 772;
Best Local Similarity 57.1%; Pred. No. 0.19;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

GN CDH8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_Taxid=9606;
OX [1]
RP SEQUENCE FROM N.A.
RX PubMed=10861224;
RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
RT "Identification of three human type-II classic cadherins and frequent
RT heterophilic interactions between different subclasses of type-II
RT classic cadherins.";
RL Biochem. J. 349:159-167(2000).
RN [2]
RP SEQUENCE OF 7-799 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95073006; PubMed=7982033;
RA Tanihara H., Sano K., Helmark R.L., St John T., Suzuki S.;
RT "Cloning of five human cadherins clarifies characteristic features of
RT cadherin extracellular domain and provides further evidence for two
RT structurally different types of cadherin.";
RL Cell Adhes. Commun. 2:15-26(1994).
RN [3]
RP SEQUENCE OF 294-799 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=91283540; PubMed=2059658;
RA Suzuki S., Sano K., Tanihara H.;
RT "Diversity of the cadherin family: evidence for eight new cadherins
RT in nervous tissue.";
RL Cell Regul. 2:261-270(1991).
RN -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Mainly expressed in brain. Found in certain
CC nerve cell lines, such as retinoblasts, glioma cells and
CC neuroblasts.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB035305; BAA87419.1; -.
DR EMBL; L34060; AAA35628.1; ALT_INIT.
DR HSSP; P15116; INCU.
DR Genew; HGNC:1767; CDH8.
DR MIM; 603008; -.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR00233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SMO0112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS50268; CADHERIN_2; 5.
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 61 POTENTIAL.
FT CHAIN 62 799 CADHERIN-8.
FT DOMAIN 62 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 642 POTENTIAL.
FT DOMAIN 643 799 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 642 167 CADHERIN_1.
FT DOMAIN 168 276 CADHERIN_2.
FT DOMAIN 277 391 CADHERIN_3.

```



```

FT DOMAIN 392 494 CADHERIN 4.
FT DOMAIN 495 616
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 37 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 355 355 V -> D (IN REF. 2 AND 3).
FT CONFLICT 647 647 H -> HQ (IN REF. 2 AND 3).
SQ SEQUENCE 799 AA; 88253 MW; 9B119B86039C6A0A CRC64;

Query Match 53.3%; Score 49; DB 1; Length 799;
Best Local Similarity 63.6%; Pred. No. 3.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 W1WNOQHIDE12
Db 63 W1WNOQFVLEE 73

RESULT 6
CAD8_MOUSE STANDARD; PRT; 799 AA.
ID CAD8_MOUSE
AC P97291;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-8 precursor.
GN CDH8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster / NIH;
RX MEDLINE=97174321; PubMed=9022055;
RA Korematsu K., Redies C.;
RT "Restricted expression of cadherin-8 in segmental and functional
subdivisions of the embryonic mouse brain.";
RL Dev. Dyn. 208:178-189(1997).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
They preferentially interact with themselves in a homophilic
manner in connecting cells; cadherins may thus contribute to the
sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95600; CAA64857.1; -
DR HSSP; P15116; INCU.
DR MGD; MGI:107434; Cdh8.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00323; CADHERIN_1; 3.
DR PROSITE; PS00368; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 61 POTENTIAL.
FT CHAIN 62 799 CADHERIN-8.

```

```

FT DOMAIN 62 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 642 POTENTIAL.
FT DOMAIN 643 799 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 62 167 CADHERIN 1.
FT DOMAIN 168 276 CADHERIN 2.
FT DOMAIN 277 391 CADHERIN 3.
FT DOMAIN 392 494 CADHERIN 4.
FT DOMAIN 495 616 CADHERIN 5.
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 799 AA; 88200 MW; 0E35FEAD563C7F76 CRC64;

Query Match 53.3%; Score 49; DB 1; Length 799;
Best Local Similarity 63.6%; Pred. No. 3.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 W1WNOQHIDE12
Db 63 W1WNOQFVLEE 73

RESULT 7
CAD8_RAT STANDARD; PRT; 799 AA.
ID CAD8_RAT
AC O54800; O54801;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cadherin-8 precursor.
GN CDH8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=98190518; PubMed=9521872;
RA Kido M., Obara S., Tanihara H., Koehelle J.M., Seldin M.F.,
RA Takekoshi S., Suzuki S.T.;
RT "Molecular properties and chromosomal location of cadherin-8.";
RL Genomics 48:186-194(1998).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
They preferentially interact with themselves in a homophilic
manner in connecting cells; cadherins may thus contribute to the
sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=O54800-1; Sequence=Displayed;
Name=2;
IsoId=O54800-2; Sequence=VSP 000638, VSP 000639;
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB010436; BAA24452.1; -
DR EMBL; AB010437; BAA24453.1; -
DR HSSP; P15116; INCU.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.

```

[illegible]

Query Match	53.3%	Score 49	DB 1	Length 799
Best Local Similarity	63.6%	Pred. No.	3.8	
Matches	7	Conservative	2	Mismatches 0; Gaps 0;

```
QY      2 W I W N Q M H I D E E 12
        | : | | | | : | |
Db      63 W W W N Q M F V L E E 73
```

## RESULT 8

CHICKEN	STANDARD;	PRT;	792 AA.
ID_CADB_CHICK			

```
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Cadherin-11 precursor.
```

05 Gallus gallus (Chicken).  
0C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
0C Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OX NCBI\_TaxID=9031;

RP SEQUENCE FROM N.A.  
CO. SMITH white Tobacco:

RA Sirakian=White Leghorn;  
RC Wei J., Dong X.R., Topouzis S., Zimmer W.E., Broders F., Thiery J.P.,  
RA Koteliansky V., Majesky M.W.;  
RA Koteliansky V., Majesky M.W.;

RT "Molecular cloning of chick calseverin II and its expression during smooth muscle differentiation and formation of the tunica media."; submitted (MD-1099) to the EMBL/GenBank/DBI databases.

CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins  
CC They preferentially interact with themselves in a homophilic

CC manner in connecting cells; cadherins may thus contribute to the sorting of heterogeneous cell types.

```
CC - SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC - SIMILARITY: Contains 5 cadherin domains.
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC of the Swiss Institute of Bioinformatics and the EMBL Outstation at  
CC the European Molecular Biology Laboratory, Heidelberg, Germany.

CC between the Swiss Institute of Bioinformatics and the EMBL Collaboration  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as the content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@rsh.ch](mailto:license@rsh.ch).

Q0	SEQUENCE	792 AA; 87572 MW; 333488C6686731AB CRC64;
FT	DOMAIN	384 486
FT	DOMAIN	487 608
FT	CARBOHYD	455 455
FT	CARBOHYD	536 536
FT	CARBOHYD	594 594
DR	PROSITE; P500268; cadherin; 5.	
DR	Pfam; PF01049; Cadherin_C term; 1.	
DR	PRINTS; PR00205; CADHERIN_	
DR	SMART; SM00112; Ca; 5.	
DR	PROSITE; P500232; CADHERIN_1; 3.	
DR	PROSITE; P50268; CADHERIN_2; 5.	
KW	Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat; signal.	
FT	SIGNAL	1 22
FT	PROPEP	23 53
FT	CHAIN	54 792
FT	DOMAIN	54 613
FT	TRANSMEM	614 613
FT	DOMAIN	635 792
FT	DOMAIN	54 159
FT	DOMAIN	160 268
FT	DOMAIN	269 383
FT	DOMAIN	384 486
FT	DOMAIN	487 608
FT	CARBOHYD	455 455
FT	CARBOHYD	536 536
FT	CARBOHYD	594 594
Q0	SEQUENCE	792 AA; 87572 MW; 333488C6686731AB CRC64;

Query Match	48.98%	Score	45	DB	1	Length	152
Best Local Similarity	54.5%	Pred. No.	16				
Matches	6	Conservative	2	Mismatches	3	Indels	0
						Gaps	0

QY	2	WIMNQMHIDEE	12
		: : : :	
Db	55	WWMNQFFVIEE	65

## RESULTS

ID	CADB HUMAN	STANDARD;	PRT;	796 AA
NC	REF387.	015065.	091093.	091094:

DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last amotetion update)  
DE Cadherin-11 precursor (Osteoblast-cadherin) (OSF-4)  
DE cont1

OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo  
NCBI TaxID: 9606.

RN [1]  
PD SENTENCE FROM N.A. (ISOFORM 1).

RC	TISSUE=Brain;
RX	MEDLINE=95073006; PubMed=7982033;

RA Tanihara H., Sano K., Heimark R.L., St John T., Suzuki S.;  
RT "Cloning of five human cadherins clarifies characteristic features of  
PT cadherin extracellular domain and provides further evidence for two

RT structurally different types of cadherin.";  
Cell Adhes. Commun. 2:15-26(1994).  
RL

RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RC TISSUE=Osteosarcoma;  
MEDLINE=94216322; PubMed=8163513;  
RX Osteosarcoma

RA Okazaki M., Takeshita S., Kawai S., Akuno K., Isubama M.,  
RA Kudo A., Amann E.;  
PM Molecular cloning and characterization of OB-cadherin, a new member

RT of cadherin family expressed in osteoblasts.";  
J Biol Chem. 269:12092-12098(1994).

RP SEQUENCE OF 269-796 FROM N.A. (ISOFORM 1).

RC TISSUE=Petal brain;  
RX MEDLINE=91283540; PubMed=2059658;

RA Suzuki S., Sano K., Tanihara H.;  
RT "Diversity of the cadherin family: evidence for eight new cadherins



RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.B., Jones S.U.W., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN  
 RP DEVELOPMENTAL STAGE.  
 RC STRAIN=CS7BL/6; TISSUE=Testis;  
 RX MEDLINE=97033837; PubMed=8879495;  
 RA Munro S.B., Blaschuk O.W.;  
 RT "A comprehensive survey of the cadherins expressed in the testes of  
 RT fetal, immature, and adult mice utilizing the polymerase chain  
 RT reaction.";  
 RL Biol. Reprod. 55:822-827(1996).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Selectively expressed in osteoblastic cell  
 CC lines, precursor cell lines of osteoblasts, and primary  
 CC osteoblastic cells from calvaria, as well as in lung, testis, and  
 CC brain tissues at low levels.  
 CC -1- DEVELOPMENTAL STAGE: In the testis, expression is highest in fetal  
 CC gonad and decreases 8-fold in newborn.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: X77557; CAA54674.1; -;  
 CC EMBL: D31963; BAA06730.1; -;  
 CC EMBL: D21253; BAA04797.1; -;  
 CC EMBL: BC046314; AAB46314.1; -;  
 CC PIR: A53584; A53584.  
 CC PIR: I48277; I48277.  
 CC PIR: I49556; I49556.  
 CC HSSP: P15116; INCI.  
 CC MGD: MGI:99217; Cdh11.  
 CC GO: GO:0005737; C:cytoplasm; IDA.  
 CC GO: GO:0005886; C:plasma membrane; IDA.  
 CC InterPro: IPR002126; Cadherin.  
 CC InterPro: IPR000233; Cadherin\_C\_term.  
 CC Pfam: PF00028; cadherin\_5.  
 CC Pfam: PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS: PR00205; CADHERIN.  
 CC SMART: SM00112; CA; 5.  
 CC PROSITE: PS00232; CADHERIN\_1; 3.  
 CC PROSITE: PS50268; CADHERIN\_2; 5.  
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 24  
 FT PROPEP 25 53  
 FT CHAIN 54 796  
 FT DOMAIN 54 617  
 FT TRANSMEM 618 640  
 FT SIGNAL 641 640  
 FT DOMAIN 641 796  
 FT DOMAIN 796 159  
 FT DOMAIN 160 268  
 FT DOMAIN 269 383  
 FT DOMAIN 384 486  
 FT DOMAIN 487 612  
 FT CARBOHYD 455 455  
 FT CARBOHYD 540 540  
 FT CARBOHYD 540 540  
 FT CONFLICT 462 462  
 FT CONFLICT 589 589  
 FT CONFLICT 655 655  
 FT  
 D -> N (IN REF. 2).  
 D -> N (IN REF. 2).

FT CONFLICT 751 751 V -> M (IN REF. 1).  
 FT CONFLICT 777 777 P -> Q (IN REF. 2).  
 FT CONFLICT 782 782 L -> P (IN REF. 2).  
 SQ SEQUENCE 796 AA; 88112 MW; 0D584D2641DD529 CRC64;  
 Query Match 48.9%; Score 45; DB 1; Length 796;  
 Best Local Similarity 54.5%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 2 WINNOMIDE 12  
 Db 55 WWMQFVIE 65  
 RESULT 11  
 CADM\_MOUSE  
 ID CADM\_MOUSE STANDARD; PRT; 813 AA.  
 AC Q9WTP5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cadherin-22 precursor (PB-cadherin).  
 GN CDH22.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.  
 RC STRAIN=ICR; TISSUE=Brain;  
 RX MEDLINE=99326347; PubMed=10398531;  
 RA Kitajima K., Koshimizu U., Nakamura T.;  
 RT "Expression of a novel type of classic cadherin, PB-cadherin in  
 RT developing brain and limb buds.";  
 RL Dev. Dyn. 215:206-214(1999).  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. PB-cadherins may have a role  
 CC in the morphological organization of pituitary gland and brain  
 CC tissues.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in brain. Abundant in  
 CC olfactory bulb, cerebrum, and cerebellum, less in pons, medulla,  
 CC and spinal cord. Low expression in heart. No expression in lung,  
 CC liver, spleen, kidney, testis, stomach, intestine, colon, and  
 CC placenta.  
 CC -1- DEVELOPMENTAL STAGE: Expressed at 9.5 dpc onwards. At 10.5 dpc, in  
 CC brain (telencephalic vesicles and isthmus), spinal cord and limb  
 CC buds (in the zone of polarizing activity). At 14.5 dpc, in  
 CC olfactory bulb and cerebellum.  
 CC -1- INDUCTION: Down-regulated by thyroid hormone.  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: AB019618; BAA34426.1; -;  
 CC MGD: MGI:1341843; Cdh22.  
 CC HSSP: P15116; INCI.  
 CC InterPro: IPR002126; Cadherin.  
 CC InterPro: IPR000233; Cadherin\_C\_term.  
 CC Pfam: PF00028; cadherin\_5.  
 CC Pfam: PF01049; Cadherin\_C\_term; 1.  
 CC PRINTS: PR00205; CADHERIN.  
 CC SMART: SM00112; CA; 5.  
 CC PROSITE: PS00232; CADHERIN\_1; 2.  
 CC PROSITE: PS50268; CADHERIN\_2; 5.

```
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL. 1 33 POTENTIAL.
FT CHAIN 34 813 CADHERIN-22.
FT DOMAIN 33 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 642 POTENTIAL.
FT DOMAIN 643 813 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 61 165 CADHERIN 1.
FT DOMAIN 166 274 CADHERIN 2.
FT DOMAIN 392 495 CADHERIN 3.
FT DOMAIN 392 495 CADHERIN 4.
FT DOMAIN 496 613 CADHERIN 5.
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 813 AA; 88021 MW; 5510F9848D976567 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 813;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMNQHIDE 12
Db 61 WWMNQFFVVE 71

RESULT 12
CADM_RAT STANDARD: PRT; 813 AA.
AC Q63J15; Q63561;
DC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cadherin-22 precursor (PB-cadherin).
OS CDH22.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RC STRAIN=WiStar; TISSUE=Brain, and Pituitary;
MEDLINE=9621232; PubMed=8626716;
RA Sugimoto K., Honda S., Yamamoto T., Ueki T., Monden W., Kaji A.,
RA Matsunoto K., Nakamura T.,
RT "Molecular cloning and characterization of a newly identified member
RT of the cadherin family, PB-cadherin."
RL J. Biol. Chem. 271:11548-11556(1996).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. PB-cadherins may have a role
CC in the morphological organization of pituitary gland and brain
CC tissues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q63J15-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q63J15-2; Sequence=VSP_000643, VSP_000644;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Strongly expressed in the pituitary gland and
CC the brain (in the inner granular and glomerular layers of the
CC olfactory bulb, anterior olfactory nucleus, primary olfactory
CC cortex, Purkinje cell layer of cerebellum, and pineal gland). Low
CC expression in lung and heart. No expression in submandibular
CC gland, thymus, liver, spleen, adrenal, and kidney.
CC -!- DEVELOPMENTAL STAGE: Expressed strongly in fetal brain.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D83348; BAA11894.1; -.
DR EMBL; D83349; BAA11895.1; -.
DR HSSP; P15116; INCU.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal; Alternative splicing.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 813 CADHERIN-22.
FT DOMAIN 33 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 642 POTENTIAL.
FT DOMAIN 643 813 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 61 165 CADHERIN 1.
FT DOMAIN 166 274 CADHERIN 2.
FT DOMAIN 275 391 CADHERIN 3.
FT DOMAIN 392 495 CADHERIN 4.
FT DOMAIN 496 613 CADHERIN 5.
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 668 694 NDEGGEQPTFAVSALRSIVDFGL -> ITGSGHRTGS
FT /FTId=VSP_000643.
FT VARSPLIC 695 813 Missing (in isoform 2).
FT /FTId=VSP_000644.
SQ SEQUENCE 813 AA; 87978 MW; 30BEA60B5D2D467B CRC64;

Query Match 48.9%; Score 45; DB 1; Length 813;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMNQHIDE 12
Db 61 WWMNQFFVVE 71

RESULT 13
CADO_HUMAN STANDARD: PRT; 819 AA.
AC O86UF0; O86UP1; Q9N784;
DC 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cadherin-24 precursor (UMQ2834/PRO34009).
GN CDH24 OR CDH11L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND INTERACTIONS WITH
RP CATECHINS.
RC MEDLINE=22753805; PubMed=12734196;
RA Kafatisz B.J., Nisnam M.T., Wheelock M.J., Johnson K.R.,
RT "Characterization of cadherin-24, a novel alternatively spliced type
RT II cadherin."
RL J. Biol. Chem. 278:27513-27519 (2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC MEDLINE=22887296; PubMed=12975309;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
```

RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Baton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simons L., Singh J., Smith V., Stinson J., Vagstad A.,  
 RA Vanden R., Watanabe C., Wiand D., Woods K., Xie M.-H., Yansura D.,  
 RA Yi S., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment.";  
 RL Genome Res. 13:2265-2270(2003).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Testis;  
 RA Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Miemann S.;  
 RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. Cadherin-24 mediate strong  
 CC cell-cell adhesion.  
 CC -1- SUBUNIT: Associates with alpha-, beta- and delta-catenins.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name=1; Synonyms=Long form;  
 CC IsoId=Q86UP0-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Short form;  
 CC IsoId=Q86UP0-2; Sequence=VSP\_008717;  
 CC Name=3;  
 CC IsoId=Q86UP0-3; Sequence=VSP\_008718, VSP\_008719;  
 CC Note=No experimental confirmation available;  
 CC -1- SIMILARITY: Contains 5 cadherin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AY260900; AAP20590.1; -;  
 DR EMBL; AY260901; AAP20591.1; -;  
 DR EMBL; AY358199; AAQ88566.1; -;  
 DR EMBL; AL137477; CAB70758.1; -;  
 DR PIR; T46418; T46418.  
 DR GeneW; HGNC:14265; CDH24.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR002033; Cadherin\_C-term.  
 DR Pfam; PF000283; cadherin; 5.  
 DR Pfam; PF01049; Cadherin\_C-term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 4.  
 DR PROSITE; PS00232; CADHERIN 1; 2.  
 DR PROSITE; PS0268; CADHERIN 2; 5.  
 DR Cell address: Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KM Signal; Multigene family; Alternative splicing.  
 FT SIGNAL 1  
 FT PROPEP 19  
 FT CHAIN 21  
 FT DOMAIN 45  
 FT TRANSMEM 642  
 FT DOMAIN 663  
 FT DOMAIN 151  
 FT DOMAIN 259  
 FT DOMAIN 260  
 FT DOMAIN 375  
 FT DOMAIN 517  
 FT CARBOHYD 446  
 FT CARBOHYD 548  
 FT CARBOHYD 563

FT VARSPPLIC 455 492 Missing (in isoform 2).  
 FT VARSPPLIC 1 427 /Frid=VSP\_008717.  
 FT VARSPPLIC 428 492 Missing (in isoform 3).  
 FT VARSPPLIC 428 492 /Frid=VSP\_008718.  
 FT VARSPPLIC 428 492 EGTHTRAPIDRRARARNHNTVATLQSGNGRGVPLL  
 FT VARSPPLIC 428 492 VAEWSAPAPQPSVGSAGVIFQ -> NNTVCTWYSIHS  
 FT VARSPPLIC 428 492 ATLFSTCTIAYRPFCLMAYASCGIHAAHPHLRVNCVC  
 FT VARSPPLIC 428 492 VWRVCFGLPSS (in isoform 3).  
 FT VARSPPLIC 428 492 /Frid=VSP\_008719.  
 FT SEQUENCE 819 AA; 87751 MM; 9083034F18BA7E4A CRC64;  
 FT SEQUENCE 819 AA; 87751 MM; 9083034F18BA7E4A CRC64;  
 QY 2 WINNOMHIDE 12  
 Db 46 WWMNQFVIBE 56  
 Query Match 48.9%; Score 45; DB 1; Length 819;  
 Best Local Similarity 54.5%; Pred. No. 17;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 RESULT 14  
 CDM\_HUMAN  
 ID CDM\_HUMAN STANDARD; PRT; 828 AA.  
 AC Q9UJ99; O43205;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cadherin-22 precursor.  
 GN CDH22 OR C20ORP25.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.B., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall J.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehaeslalo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie J., McIlroy K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symamore N.C., Taylor R., Tee L., Thomas D.W., Thorne A.,  
 RA Tracey A., Tromans A.C., Vaundin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmington L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [2]  
 RP SEQUENCE OF 449-828 FROM N.A.  
 RC TISSUE=Brain;  
 RA Yu W., Sarajedini J., Gibbs R.A.;  
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.  
 CC They preferentially interact with themselves in a homophilic  
 CC manner in connecting cells; cadherins may thus contribute to the  
 CC sorting of heterogeneous cell types. PB-cadherins may have a role  
 CC in the morphological organization of pituitary gland and brain

```

CC tissues (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL031687; CAB51587.2; -.
DR EMBL; AF035300; AAB88183.1; -.
DR Genew; HGNC:13251; CDH22.
DR HSSP; P15116; 1NCJ.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS00268; CADHERIN_2; 5.
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 828 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 625 624 POTENTIAL.
FT DOMAIN 625 645 POTENTIAL.
FT DOMAIN 646 828 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 169 168 CADHERIN_1.
FT DOMAIN 169 277 CADHERIN_2.
FT DOMAIN 278 394 CADHERIN_3.
FT DOMAIN 395 439 CADHERIN_4.
FT DOMAIN 439 498 CADHERIN_5.
FT DOMAIN 498 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 828 AA; 89091 MW; 52077BD1F624DCA CRC64;

Query Match 48.9%; Score 45; DB 1; Length 828;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WWMNQMHIDE 12
Db 64 WWMNQFVLEE 74

RESULT 15
CAD7 CHICK STANDARD; PRT; 785 AA.
AC Q90763.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cadherin-7 precursor.
GN CDH7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Brain;
RX MEDLINE=95309115; PubMed=7540531;
RA Nakagawa S., Takeichi M.;
RT "Neural crest cell-cell adhesion controlled by sequential and
RT subpopulation-specific expression of novel cadherins.";
RL Development 121:1321-1332(1995).
-1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.

```

```

CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D42150; BAA07721.1; -.
DR PIR; I50180; I50180.
DR HSSP; P15116; 1NCH.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS00268; CADHERIN_2; 5.
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 27 POTENTIAL.
FT PROPEP 28 47 POTENTIAL.
FT CHAIN 48 785 CADHERIN-7.
FT DOMAIN 48 607 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 608 628 POTENTIAL.
FT DOMAIN 629 785 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 154 153 CADHERIN_1.
FT DOMAIN 154 262 CADHERIN_2.
FT DOMAIN 263 377 CADHERIN_3.
FT DOMAIN 378 482 CADHERIN_4.
FT DOMAIN 482 599 CADHERIN_5.
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 785 AA; 87171 MW; 895B06D8141B34D4 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 785;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WWMNQMHIDE 12
Db 49 WWMNQFVLEE 59

Search completed: July 29, 2004, 11:35:23
Job time : 7.66667 secs

```

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2004, 11:31:25 ; Search time 10.3333 Seconds  
(without alignments)  
139.633 Million cell updates/sec

Title: US-10-040-128-3  
Perfect score: 92  
Sequence: 1 DWIMQMHIDEKNT 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	92	100.0	784	1 IJHUC5	cadherin 5 precursor
2	49	53.3	793	2 D38992	cadherin 8 - human
3	48	52.2	524	2 T20872	hypothetical prote
4	45	48.9	796	2 A38992	cadherin 11 precu
5	45	48.9	796	2 I48277	cadherin-11 - mous
6	45	48.9	796	2 I49556	cadherin-11 - mous
7	45	48.9	796	2 A53584	OB-cadherin precu
8	44.5	48.4	963	2 T09911	probable serine/th
9	44	47.8	500	2 B89962	hypothetical prote
10	44	47.8	785	2 I50180	cadherin-7 - chick
11	44	47.8	790	2 I51638	F-cadherin - Afric
12	44	47.8	790	2 G02678	cadherin-14 - huma
13	44	47.8	794	2 I59372	cadherin 12 - huma
14	43	46.7	435	2 C84941	histidinol dehydro
15	42.5	46.2	388	1 D70006	conserved hypothet
16	42	45.7	82	2 T46510	hypothetical prote
17	42	45.7	296	2 F95380	probable hydrolase
18	42	45.7	315	2 AG3535	dipeptide transpor
19	42	45.7	427	2 AF0590	citrate synthase l
20	42	45.7	529	2 T21575	hypothetical prote
21	42	45.7	907	2 E8692	probable wall-asso
22	41.5	45.1	392	2 S55488	probable NADPH deh
23	41.5	45.1	1032	2 T43257	beta-1,3 exoglucan
24	41	44.6	77	2 A11306	hypothetical prote
25	41	44.6	314	2 S47696	nucleic transport s
26	41	44.6	314	2 H91171	transport of nucle
27	41	44.6	314	2 H86017	transport of nucle
28	41	44.6	431	2 B82119	citrate synthase V
29	41	44.6	519	2 T33616	hypothetical prote

30	41	44.6	532	2 G90607	hypothetical prote
31	41	44.6	619	2 B97668	hypothetical prote
32	41	44.6	619	2 AG2892	hypothetical prote
33	41	44.6	712	2 T16338	hypothetical prote
34	41	44.6	770	2 S11161	anthranilate synth
35	41	44.6	789	2 I52701	K-cadherin - rat
36	41	44.6	790	2 I37016	cadherin-6 - human
37	41	44.6	790	2 I50178	cadherin-6B - chlc
38	41	44.6	2241	2 T20971	hypothetical prote
39	41	44.6	2261	2 T20978	hypothetical prote
40	41	44.6	6831	2 A88852	protein unc-22 (lm
41	41	44.6	6839	2 S57242	twistchin [similar
42	41	44.6	7160	2 T27935	hypothetical prote
43	40	43.5	77	2 A11678	hypothetical prote
44	40	43.5	120	2 C72425	hypothetical prote
45	40	43.5	200	1 G64106	3-isopropylmaleate

ALIGNMENTS

RESULT 1  
IJHUC5  
cadherin 5 precursor - human  
N/Alternate names: 7B4 antigen; cadherin, endothelial-specific; VE-cadherin  
C/Species: Homo sapiens (man)  
C/Date: 30-Jun-1993 #sequence revision 13-Sep-1996 #text change 22-Jun-1999  
C/Accession: S49893; S24305; A43418  
R/Breviario, F.; Cavada, L.; Corada, M.; Martin-Padura, I.; Golay, J.; Introna, M.; Lamp  
submitted to the EMBL Data Library, June 1994  
A/Description: Molecular and functional properties of VE-cadherin (7B4/cadherin-5) a nov  
A/Reference number: S49893  
A/Accession: S49893  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-784 <BRE>  
A/Cross-references: EMBL:X79981; NID:G599833; PIDN:CA45306.1; PID:G599834  
R/Suzuki, S.; Sano, K.; Tanihara, H.  
Cell Regul. 2, 261-270, 1991  
A/Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous t  
A/Reference number: S24305; MUID:91283540; PMID:2059658  
A/Accession: S24305  
A/Molecule type: mRNA  
A/Residues: 5-516, 'I', 518-784 <SUZ>  
A/Cross-references: EMBL:X59796; NID:G639976; PIDN:CA442468.1; PID:G29593  
R/Lampugnani, M.G.; Resnati, M.; Ralteri, M.; Pigotti, R.; Pissacane, A.; Houen, G.; Ruco,  
J. Cell Biol. 118, 1511-1522, 1992  
A/Title: A novel endothelial-specific membrane protein is a marker of cell-cell contacts  
A/Reference number: A43418; MUID:92394977; PMID:1522121  
A/Accession: A43418  
A/Molecule type: protein  
A/Residues: 48-60, 'X', 62, 'X', 64, 108-116, 'X', 118-123, 237-238, 'X', 240, 'X', 242-252, 'X', 254-  
A/Experimental source: cultured endothelial cells  
A/Note: sequence extracted from NCBI Backbone (NCBI:113040, NCBI:113045, NCBI:113047,  
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought t  
C/Genetics:  
A/Gene: GDB:CDH5  
A/Cross-references: GDB:134230; OMIM:601120  
A/Map position: 16q22.1-16q22.1  
C/Superfamily: cadherin; cadherin repeat homology  
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro  
F/1-25/Domain: signal sequence #status predicted <SIG>  
F/26-47/Domain: propeptide #status predicted <PRO>  
F/48-784/Product: cadherin 5 #status predicted <MAT>  
F/48-593/Domain: extracellular #status predicted <EXT>  
F/150-151/Domain: cadherin repeat homology <CR1>  
F/154-258/Domain: cadherin repeat homology <CR2>  
F/261-372/Domain: cadherin repeat homology <CR3>  
F/375-479/Domain: cadherin repeat homology <CR4>  
F/481-587/Domain: cadherin repeat homology <CR5>  
F/594-620/Domain: transmembrane #status predicted <TM>  
F/621-784/Domain: intracellular #status predicted <INT>  
F/736-753/Region: serine-rich

F:61,112,157,362,442,523,535/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 100.0%; Score 92; DB 1; Length 784;  
Best Local Similarity 100.0%; Pred. No. 5.8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWIMQMHIDEKNT 15  
|||||  
48 DWIMQMHIDEKNT 62

## RESULT 2

D38992  
C:Species: Homo sapiens (man)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jan-2000  
C:Accession: D38992  
R:Suzuki, S.; Sano, K.; Tanihara, H.  
Cell Regul. 2, 261-270, 1991  
A:Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous ti  
A:Reference number: S24305; MUID:91283540; PMID:2059658  
A:Accession: D38992  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-793 <SUZ>  
A:Cross-references: GB:J34060; NID:G506411; PIDN:AAA5628.1; PID:G506412  
C:Genetics:  
A:Gene: GDB:CDH8  
A:Cross-references: GDB:5822911  
C:Superfamily: cadherin; cadherin repeat homology  
C:Keywords: calcium binding; cell adhesion; duplication  
F:163-269/Domain: cadherin repeat homology <CDH>

Query Match 53.3%; Score 49; DB 2; Length 793;  
Best Local Similarity 63.6%; Pred. No. 6.2;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WIMQMHIDE 12  
|||||  
56 WIMQMHIDE 66

## RESULT 3

T20872  
hypothetical protein F26H9.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T20872; T21443  
R:Sim, M.  
submitted to the EMBL Data Library, June 1996  
A:Reference number: Z19337  
A:Accession: T20872  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-524 <WIL>  
A:Cross-references: EMBL:Z75535; PIDN:CAA9829.1; GSPDB:GN00019; CESP:F26H9.1  
A:Experimental source: clone F14B4  
R:Baynes, C.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19422  
A:Accession: T21443  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-524 <W12>  
A:Experimental source: EMBL:Z81516; PIDN:CA804206.1; GSPDB:GN00019; CESP:F26H9.1  
C:Genetics:  
A:Gene: CESP:F26H9.1  
A:Map position: 1  
A:Insertions: 49/2; 99/1; 136/1; 186/1; 228/3; 287/3; 346/1; 403/1; 449/3

Query Match 52.2%; Score 48; DB 2; Length 524;  
Best Local Similarity 53.8%; Pred. No. 5.7;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMQMHIDEKNT 14  
|||||  
486 WIMQMHIDEKNT 498

## RESULT 4

A38992  
cadherin 11 precursor - human  
N:Alternate names: OB-cadherin, osteoblast  
C:Species: Homo sapiens (man)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 20-Aug-1999  
C:Accession: A38992  
R:Suzuki, S.; Sano, K.; Tanihara, H.  
Cell Regul. 2, 261-270, 1991  
A:Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous ti  
A:Reference number: S24305; MUID:91283540; PMID:2059658  
A:Accession: A38992  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-796 <SUZ>  
A:Cross-references: GB:J34056; NID:G506403; PIDN:AAA5622.1; PID:G506404  
C:Genetics:  
A:Gene: GDB:CDH11; OB  
A:Cross-references: GDB:512891; OMIM:600023  
A:Map position: 16q22.1-16q22.1  
C:Superfamily: cadherin; cadherin repeat homology  
C:Keywords: calcium binding; cell adhesion; duplication  
F:156-159/Domain: cadherin repeat homology <CR1>  
F:162-268/Domain: cadherin repeat homology <CR2>  
F:271-383/Domain: cadherin repeat homology <CR3>  
F:386-488/Domain: cadherin repeat homology <CR4>

Query Match 48.9%; Score 45; DB 2; Length 796;  
Best Local Similarity 54.5%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMQMHIDE 12  
|||||  
55 WIMQMHIDE 65

## RESULT 5

I48277  
cadherin-11 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jan-2000  
C:Accession: I48277  
R:Hoffmann, I.; Bailling, R.  
Dev. Biol. 169, 337-346, 1995  
A:Title: Cloning and expression analysis of a novel mesodermally expressed cadherin.  
A:Reference number: I48277; MUID:95269886; PMID:7750649  
A:Accession: I48277  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-796 <RES>  
A:Cross-references: EMBL:X77557; NID:G642796; PIDN:CAA54674.1; PID:G666071  
C:Genetics:  
A:Gene: cad-11  
C:Superfamily: cadherin; cadherin repeat homology <CDH>  
F:56-159/Domain: cadherin repeat homology <CDH>

Query Match 48.3%; Score 45; DB 2; Length 796;  
Best Local Similarity 54.5%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WIMQMHIDE 12  
|||||  
55 WIMQMHIDE 65

## RESULT 6



C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 20-Aug-1999  
 C/Accession: I51638; S55391  
 R/Species: A.; Johnson, E.; Kintner, C.  
 M/I: Cell Neurosci. 6, 199-212, 1995  
 A/Title: Xenopus F-cadherin, a novel member of the cadherin family of cell adhesion mole  
 A/Reference number: I51638; PMID:96039533; PMID:7496627  
 A/Accession: I51638  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-790 <ESP>  
 A/Cross-references: EMBL:X85330; NID:9854634; PIDN:CAA59679.1; PID:9854635  
 C/Superfamily: cadherin; cadherin repeat homology  
 F:161-267/Domain: cadherin repeat homology <CR2>

Query Match 47.8%; Score 44; DB 2; Length 790;  
 Best Local Similarity 54.5%; Pred. No. 41;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WNWQMHIDE 12  
 Db 54 WWNQFVLEE 64

RESULT 12  
 G02678  
 cadherin-14 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 21-Jan-2000  
 C/Accession: G02678  
 R/Shibata, T.; Shimoyama, Y.; Gotch, M.; Hirohashi, S.  
 submitted to the EMBL Data Library, May 1996  
 A/Reference number: H01584  
 A/Accession: G02678  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-790 <SHI>  
 A/Cross-references: EMBL:U59325; NID:G1389852; PIDN:AA802933.1; PID:G1389853  
 C/Superfamily: cadherin; cadherin repeat homology  
 F:162-268/Domain: cadherin repeat homology <CDH>

Query Match 47.8%; Score 44; DB 2; Length 790;  
 Best Local Similarity 54.5%; Pred. No. 41;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WNWQMHIDE 12  
 Db 55 WWNQFVLEE 65

RESULT 13  
 I59372  
 cadherin 12 - human  
 N/Alternate names: Br-cadherin  
 C/Species: Homo sapiens (man)  
 C/Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 20-Aug-1999  
 C/Accession: I59372  
 R/Selley, S.; Bruno, S.; Scharf, U.M.; Wang, C.H.; Vitale, E.; Gilliam, T.C.; Kunkel, L.M.  
 Proc. Natl. Acad. Sci. U.S.A. 92, 3702-3706, 1995  
 A/Title: Expressed cadherin pseudogenes are localized to the critical region of the spri  
 A/Reference number: I59372; PMID:95249541; PMID:7731968  
 A/Accession: I59372  
 A/Status: nucleic acid sequence not shown; translation not shown; translated from GB/EM  
 A/Molecule type: mRNA  
 A/Residues: 1-794 <RNS>  
 A/Cross-references: GB:U3477; NID:G793942; PIDN:AA848539.1; PID:G793943  
 C/Genetics:  
 A/Gene: GDB:CDH12  
 A/Cross-references: GDB:596324  
 A/Map position: 5p13-5p14  
 C/Superfamily: cadherin; cadherin repeat homology  
 C/Keywords: calcium binding; cell adhesion; duplication; transmembrane protein  
 F:57-160/Domain: cadherin repeat homology <CR1>  
 F:163-269/Domain: cadherin repeat homology <CR2>

F:272-384/Domain: cadherin repeat homology <CR3>  
 F:387-489/Domain: cadherin repeat homology <CR4>  
 F:491-601/Domain: cadherin repeat homology <CR5>  
 F:610-637/Domain: transmembrane #status predicted <TM>  
 F:638-794/Domain: intracellular #status predicted <IN>

Query Match 47.8%; Score 44; DB 2; Length 794;  
 Best Local Similarity 54.5%; Pred. No. 41;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 WNWQMHIDE 12  
 Db 56 WWNQFVLEE 66

RESULT 14  
 C84941  
 histidinol dehydrogenase (BC 1.1.1.23) [imported] - Buchnera sp. (strain APS)  
 C/Species: Buchnera sp.  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C/Accession: C84941  
 R/Shigenobu, S.; Matanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
 Nature 407, 81-86, 2000  
 A/Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A  
 A/Reference number: A84930; PMID:20445173; PMID:10993077  
 A/Accession: C84941  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-435 <STO>  
 A/Cross-references: GB:AP000398; GSPDB:IGN00144  
 A/Experimental source: strain APS  
 C/Genetics:  
 A/Gene: hsd; BU100  
 C/Superfamily: histidinol dehydrogenase; histidinol dehydrogenase homology  
 C/Keywords: oxidoreductase

Query Match 46.7%; Score 43; DB 2; Length 435;  
 Best Local Similarity 60.0%; Pred. No. 31;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WNWQMHIDE 13  
 Db 10 WNKHPDECK 19

RESULT 15  
 D70006  
 conserved hypothetical protein yubA - Bacillus subtilis  
 C/Species: Bacillus subtilis  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C/Accession: D70006  
 R/Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.T.; Eutian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall  
 lech, J.; Hatwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maneel  
 Y., M.; Ogawa, K.; Ogilwa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecelle  
 Rieger, M.; Rivolt, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
 A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero  
 akeruch, M.; Tamakoshi, A.; Tanaka, T.; Trespstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
 A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A/Reference number: A69580; PMID:98044033; PMID:9384377  
 A/Accession: D70006  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-388 <KUN>  
 A/Cross-references: GB:Z99119; GB:AL009126; NID:G2635411; PIDN:CAB15094.1; PID:G2635600  
 A/Experimental source: strain 168

C:Genetics:  
A:Gene: yubA

C:Superfamily: Bacillus subtilis conserved hypothetical protein yueF

Query Match

Best Local Similarity 46.2%; Score 42.5; DB 1; Length 388;

Matches 9; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Oy 1 DMWNQMH-DEKN 14

Db 366 DMFKQSHYDEKN 380

Search completed: July 29, 2004, 11:37:50  
Job time : 10.333 secs

***This Page Blank (uspto)***